The Dissertation Committee for Jan Frederik Schlupp
certifies that this is the approved version of the following dissertation:

Essays on Firms in International Trade

Committee:

Olivier Coibion, Supervisor

Nitya Pandalai-Nayar, Co-Supervisor

Christopher Boehm

Aaron Flaaen

Essays on Firms in International Trade

by

Jan Frederik Schlupp

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

May 2021

Dedicated to my beloved parents Ingo and Andrea. Thanks for your unending support.

Und für meine Großeltern, besonders Helga and Brigitte. Ich wünschte, ihr wärt hier.

Acknowledgments

I wish to thank the multitudes of people who helped me. Thanks to Alexa, without whom the past five years would have been indescribably more empty of joy. Thanks to my siblings Nils, Leoni and Carolin for all the years of friendship. Time would fail me to tell of all the friends who made this possible, but I have to mention Cooper, Willis, Josh, Alexander, Tiemo, Andrew, Cooper, Kevin, and so many more . . .

Essays on Firms in International Trade

Publication No. ________________

Jan Frederik Schlupp, Ph.D.
The University of Texas at Austin, 2021

Supervisors: Olivier Coibion
Nitya Pandalai-Nayar

Macroeconomic phenomena are ultimately the aggregation of a large number of micro decisions. In the case of international trade, decisions at both the plant and the firm level ultimately drive trends in the aggregate data. All three chapters in this dissertation focus on extending our understanding of the decisions made by exporters. In the first chapter,"Agglomeration and Local Spillover Effects of US Exporters", I extend the study of local area information frictions in the context of exporting. Leveraging a novel dataset of establishment level exports in the US, I show for the first time in the US context that information about foreign demand affects entry and exit decisions into exporting at the county level. I extend a common model of social learning to incorporate the multi-dimensional nature of exporting choices. The second chapter,"The Local-Area Impact of Exporting", details the creation of the establishment level export dataset. It also uses the novel detailed look into regional trade to show the additional impact of the 2007-2009 collapse in trade during the Great Recession on local labor market outcomes in counties which are more exposed to trade

shocks. Partly inspired by the first chapter's observation that intermittent exporting is difficult to rationalize in standard models of trade, the third chapter, "US exporters between 1993-2017", documents how the export decisions of US exports vary over their life-cycle. Each of these chapters contributes new understanding of establishment export decisions at the micro-level, and taken together show that as data quality improves, the local environment of firms and exporters will increasingly matter for targeted macroeconomic interventions.

Table of Contents

List of Tables

List of Figures

Chapter 1

Agglomeration and Local Spillover Effects of US Exporters

1.1 Introduction

A look below the seemingly stable surface of aggregate export statistics in the US reveals a surprisingly dynamic and churning situation. At the most detailed level of exporter disaggregation, the establishment level, export status is volatile on a year to year basis. In the universe of establishments in the US, 2% of establishments enter into exporting for the first time every year. However, establishments which already export do not exhibit a stable pattern of export behavior. If we consider a destination market to be the pair of the Harmonised System (HS) category product and destination country, in the same sample, establishments have a 15% chance every year of changing their export portfolio at the destination market basis. This includes both intermittently exporting in a market and overall switching of export markets. Even a more restrictive view shows significant dynamism in destination country and product export choice. Establishments change the set of products they export, keeping the set of countries exported to the same, 4% of the time. Similarly, they change the set of countries exported to, keeping the set of products exported the same, 2% of the time.

This significant instability in export activity is difficult to contextualize in standard trade models. In a Melitz (2003) model with high fixed costs of export entry, one can explain this behavior with a series of large and transitory shocks to the exporter in order to overcome the inertia inherent to a system with high initial fixed costs. However, this series of shocks does not seem ex-ante probable, and invoking a series of large shocks excludes the possibility of discovering a different mechanism which results in the observed behavior.

I propose a new explanation for these phenomena using a model of social learning in which establishments are initially uncertain about demand in export markets. Establishments learn about demand in foreign markets by observing other neighboring establishments export performance, and this changing set of beliefs about export demand induces establishments to direct their entry and switching decisions in a more dynamic fashion than expected from a full information model where establishments can ex ante correctly sort the possible markets by demand potential.

I go on to leverage a novel, confidential, establishment level trade data-set in the US to provide detailed micro-evidence consistent with this model. I show that entry, exit, and switching decisions are all significantly affected by measures of neighboring establishment export performance, and that both the precision as well as the strength of the neighbor's export performance is taken into account by the establishment when it makes decisions about its export markets.

The importance of information frictions to macroeconomic outcomes has become well established over the past few years. [1] Uncertainty about contemporaneous variables and the inherent difficulty of making accurate forecasts, as well as the costliness of information

[1] Coibion and Gorodnichenko (2015)

acquisition, affect firm level decision making in profound ways. By incorporating information frictions, macroeconomic models are able to explain a diverse set of macroeconomics puzzles. In particular, frictions in the dissemination of information are critical to explaining delayed firm responses as well as rationalizing why firms under and overreact to new information.

Studies of why firms export, on the other hand, are mostly set in either static or full information frameworks. Canonical models of international trade, such as Melitz (2003), therefore give us a good understanding of the role that productivity and fixed costs play in the export decision. The Melitz model can explain a wide array of facts in the data by modeling a large one time fixed cost to begin exporting. However, the international trade literature documents several puzzling facts which cannot be explained in the context of a full information model with a one time fixed costs.

Intermittent exporting, destination and product switching, and low survival rates of exporters are difficult to rationalize in a full information framework. However, these phenomena could arise if firms are uncertain about demand for their product in the potential foreign markets. Albornoz et al. (2012) document so called "intermittent exporters" which display a fundamental uncertainty about earning profits abroad in general. Besedeš and Prusa (2011) argue that firm uncertainty about the fixed cost to export to different markets plays an important role in the evolution of aggregate export volumes in developing economies.

This paper tackles the question of how imperfect information plays a role in the export decision process using a model of local learning in the US. This question has historically been difficult to address in the US context because the US export micro data contained in the Longitudinal Firm Trade Transactions Database (LFTTD) is recorded at the firm level, making it difficult to accurately assign trade transactions to a geographical location. This is why there are no previous studies examining county level information spillovers in the US at the establishment level.

Using a novel dataset on establishment level exports gives a first potential analysis of this channel. This novel dataset merges several US Census micro-datasets in order to assign export transactions to a specific establishment within a firm. Using this detailed administrative panel data I apply a model consistent regression specification with rich fixed effects to isolate the effects of the local transmission of information known as export spillovers.

Evidence for the existence of spillovers in the US is not yet a settled debate. Seminal work by Bernard and Jensen (2004), during their investigation into why some firms export, tangentially raises the question of whether or not local exporting activity induces establishments to begin exporting themselves. Bernard and Jensen (2004) test for the existence of geographic spillover effects both at the state level, and at the industry level outside of the state of the exporting firm. In this context, they find that the effects of spillovers are either insignificant, or negative, suggesting that existing exporting behavior actually inhibits export entry. By narrowing the geographic scope I may be able to detect spillover effects that operate at a short geographic distance that are washed out in a statewide analysis.

I set up a model of local social learning, in the context of which I show that spillover effects are an important driver of establishment level export decisions in the US. This model is motivated by the Fernandes and Tang (2014) model of social learning, who study local learning by Chinese exporters. They test for firm learning from neighbors across different potential export destination countries. I extend their model to consider learning along the product dimensions as well as the destination country dimension in order to isolate and quantify about which aspects of exporting exporters learn the most. By including both the product and the country dimension, the observed set of information is neatly partitioned into three groups. Some information will match only along the potential country dimension, some information will match only along the potential product dimension, and some information will match along both dimensions.

I find that exporters are more likely to begin exporting for the first time due to the presence of exporting establishments in their county. These effects are largest for establishments which share both product category and destination country with the incumbent exporters. Exporters are also less likely to enter markets where they exit soon after entering if they have more precise information about demand in that market. From a qualitative perspective, the model does not distinguish between the effects of uncertainty about country and product demand; I find that quantitatively, establishments are more responsive to information along the product than the country dimension.

First entry into exporting decisions can be rationalized in a simple model with permanent shocks and a one time entry decision. I further add to the spillover literature by extending the model to multiple periods in order to model dynamic considerations. Establishments generally do not exhibit a stable set of export destination countries and the product mix exported. I observe a 15% chance of an establishment that exports two periods in a row changing either its choice of export destinations, products exported, or both. This churn is in part caused by the establishment's changing information set. I find that the establishment export choices are affected by information spillovers throughout its exporting life. In particular, when establishments seek to alter their portfolio of export destinations, as well as which products they export, these decisions are materially impacted by information about demand in those specific respective markets. In the dynamic context establishments still respond most strongly to information matching along both dimensions, and also react more strongly to product information rather than country information.

The importance of the extensive margin of trade, that is entry and exit into exporting by firms, has been growing over time. Goods export volume originating in the US has increased dramatically since 1990, both in absolute terms and as a proportion of US GDP. Between 1995 and 2015, the value of goods exported from the US has nearly tripled from 575 billion USD to 1.5 trillion USD.[2] Additionally, the variety of goods traded and the set of export destination countries served by the US has also grown significantly (Lincoln and McCallum, 2018). However, after controlling for firm productivity and measures of fixed costs, the remaining margin of the firm export entry decision is still not well understood. This paper extends our understanding of the extensive margin of trade on the establishment level.

The export participation choice is also a politically and policy relevant object. During the same time-frame as above, 1995 to 2015, the US trade deficit in goods grew from 174 billion USD to 760 billion USD. There are numerous export promotion agencies and policies in the US. For example, the United States Commercial Service (USCS) cites difficulties in choosing the correct export market as a key friction for domestic firms. However the literature on the effectiveness of export promotion policies still has not settled the questions of which types of EPP policies are the most useful. Given that firm productivity is likely to be a more difficult policy lever to move than information sets, the results in this paper indicate a role for US policy targeted to reduce information frictions and increase firm knowledge about

[2] U.S. Bureau of Economic Analysis, Exports of Goods

demand for their products as an effective export promotion tool.

Related Literature

This paper is situated right at the heart of the difficult question about export entry choices and their causal determinants. There is a long and unsettled history in the literature concerning the question of how local spillovers, information acquisition, and export choice interact.

Aitken et al. (1997) find in a study of Mexican firms that while the presence of multinational exporters in a region generates additional export activity by incumbent domestic non-exporters, increased activity of domestic exporters does not affect export probabilities of the same domestic firms.

Sjöholm (1999) finds the exact opposite in a study of Indonesian exporters. This study finds that district presence of FDI has no effect on entry into exporting. The paper does invoke information frictions as a primary factor affecting export decisions and contends that firms with foreign contacts and ownership are more likely to export. This effect operates along the intensive margin also, which is to say that firms with more foreign ownership are even more affected than firms with less foreign ownership share. These characteristics are clearly likely to be correlated with factors beyond information acquisition, like access international financing,

There are a number of explanations which could rationalize the above findings. It is entirely possible that information transmission operates differently in different countries. This could be due to different stages of economic development, the composition of goods exported or the composition of destination countries, or differences in the measures of FDI.

Nonetheless, this highlights the difficulty of identifying information spillovers using aggregate data.

Bernard and Jensen (2004) also find that there are no spillovers from exporters in the United States. In fact, they somewhat surprisingly find the opposite, which is that spillovers effects in the United States are negative. They find this both in terms of a general industry spillover, and a measure of same industry activity within the state. Furthermore, they find the same lack of effectiveness and statistical insignificance for export promotion policies within a state. They acknowledge the limitations of their work by pointing out the potential selection issue of focusing only on the largest firms, as well as having a fairly expansive definition of local, considering the whole state to check for spillover effects. It is quite possible that establishments with limited information capacity are unable to learn at such distance.

Furthermore, while Bernard and Jensen (2004) find no spillover effects in the United States, this finding is at odds with several empirical studies in other developed countries. The phenomenon of export spillovers and exporter agglomeration generally has been documented in countries such as France (Koenig et al. (2010)), Russia (Cassey and Schmeiser (2013)), and Spain (Castillo-Giménez et al. (2011) among others). It's important to note that while these studies vary in their definition of spillovers, they all fit into the general category

In a study using a dataset of Spanish exporters covering about 15% of Spanish exporters, Castillo-Giménez et al. (2011) find that the presence of exporters generally increases the likelihood of non-exporting incumbents to begin exporting. While their evidence for country specific spillovers is mixed, and they do not specifically investigate the mechanism by which these spillovers operate, the evidence for general local spillovers is strong.

In Koenig et al. (2010) a very detailed panel of French exporters is used to test for spillover effects from same-country same-product exporters in France. They find that firms are more likely to begin exporting to the country they observe other firms exporting to. Their findings are perfectly consistent with the findings in this paper, and the mechanisms posited in this model. Cassey and Schmeiser (2013) also posit a destination specific externality which is alleviated by destination specific spillovers in Russia.

These papers do not focus specifically on developing a model which explains these choices and mostly focus on the initial export decision. I add to this literature by examining if the same mechanisms can operate in the United States, and by examining the dynamic implications of the learning mechanism.

Until very recently, questions about the location of exporting plants in the United States have been difficult to address. Previous work relies on either aggregate data or single unit firms to properly place export transactions to their appropriate location. However, the recent research into spillovers indicates that a firm's its location and proximity to other exporting firms are important determinants of a firm's decision, not just to begin exporting but also how much to export and the duration of the export spell. However, simply showing the existence of spillover effects in the US is not enough.

It is not yet clear by which mechanisms the spillover effect works to affect establishment level export decisions. Some foundational research by Aitken et al. (1997) and Greenaway et al. (2004) looks at the link between the presence of foreign multinational exporting firms and domestic learning about export activity. Aitken et al. (1997) argue that spillover effects due to multinational exporters comes from their inherent multi-market nature. This leads them to be natural channels for transferring information about foreign markets and customers.

However, this argument discounts the information inherent in any establishment exporting decision. Greenaway et al. (2004) explore the potential for productivity spillovers coming from multinational exporters, given their likelihood of being more productive firms. However, this would not explain the dynamic decisions observed in the data. Furthermore, a preliminary version of the results in this paper did not reveal significant changes in establishment productivity.

Work by Fernandes and Tang (2014) and Choquette and Meinen (2015) generalizes those concepts to consider all firms. This paper follows in that vein of questioning, and contributes a novel insight into the spillover mechanism. In particular, I extend the Fernandes and Tang (2014) model to a more general definition of a destination market. In that paper, the firm's concept of a market is a destination-country pair. I also consider entry into exporting different products in addition to different countries. By splitting the market dimension, I gain an important additional source of variation among export decisions. This allows for a closer examination of the mechanisms which might be the cause of spillover effects. In particular, this allows me to rule out the role of supply shocks in driving export entry. I find that establishments learn about foreign demand from their local neighbours and use that information to make export decisions.

1.2 Model

The baseline version of the model which I use to investigate the possible existence of spillover effects and the potential mechanism by which this could occur is built on a 2 country Melitz (2003) model with learning about demand, extending the model of Fernandes and Tang (2014).

In a study of Chinese exporters, Fernandes and Tang build on a Jovanovic (1982) model of social learning to explain low survival rates among new exporters which they observe in their dataset. They draw inspiration from the development literature's investigation of technology adoption and note that self-experimentation can be more costly than being able to gather information about foreign markets from other firm's experimentation instead. They also argue that a story about high initial sunk costs is unlikely to explain their data.

Therefore they set up a model in which firms can learn about demand in a destination market by observing the export performance of other firms in the same city to the same destination market. An important decision in their paper is that a market is defined as a product-country pair. This means that they consider the spillovers to come from firms which match along both the country and the product dimension.

Their empirical findings are extremely promising for the learning model perspective. They find that spillover effects affect firm's entry decision as well as their initial performance in a market. They also find that the effects are stronger when more firms are around to provide the information, as well as when a firm is initially more uncertain about foreign market demand. Their findings are economically as well as statistically significant. Their findings are ambiguous in the context of firm survival, unlike what I find in my empirical results. This difference could be due to numerous factors, but remainds an interesting point of differentiation.

I extend the model to consider both the destination country and the exported product dimensions of a destination market as two separate dimensions along which establishments can learn. This additional dimension is critical for isolating the mechanism by which spillover effects operate. For example, observing certain dimension specific decisions being unresponsive to the other dimension can rule out the effects of local supply shocks.

1.2.1 Firm's problem

The household problem is purposely kept as simple as possible. On the domestic side of the economy, we observe a fixed supply of L workers with CES preferences over N goods, consuming c_i and supplying constant unit labor. The firm problem on the domestic side is also simple and closely follows Melitz (2003). Firms draw their productivity $\phi \sim F(\phi)$. Assume an iceberg style trade-cost of $\tau > 1$. A firm has a domestic and a potential foreign market for its differentiated variety of the good. Let $\rho = \frac{\sigma-1}{\sigma}$ be the standard markup charged by a firm. The firm maximizes its profits by choosing whether or not to export, and the optimal quantity of exports.

$$
\begin{aligned}
\pi &= p_i(\phi)q_i - l_i\phi \\
&= p_i(\phi)q_i - (f_d + q_i/\phi) \\
&= p_i(\phi)q_iL - (f_d + c_iL/\phi) \\
&= L\left(\frac{p_i(\phi)}{P}\right)^{-\sigma} Cp_i - \left(f_d + L\left(\frac{p_i(\phi)}{P}\right)^{-\sigma} C/\phi\right)
\end{aligned}
$$

The maximization of the above expression with respect to price yields the following optimal price expressions for domestic and foreign prices: $p_d(\phi) = \frac{1}{\rho\phi}$ and $p_x(\phi) = \frac{\tau}{\rho\phi}$

Given a domestic price, domestic revenue is a simple function of prices and productivity: $R_d(\phi) = (P\rho\phi)^{\sigma-1}R_d$

1.2.2 Uncertain demand

In the key departure from a standard Melitz model, firms face uncertain demand in the potential export location. The choice set of potential export markets is extremely large. Recent research results indicate that information acquisition is costly, even for domestic firms concerning aggregate domestic phenomena. Similarly, firm entry and exit dynamics as well as low survival rates indicate that imperfect information plays a key role in export dynamics [3]. Establishments might be uncertain about their export potential for a number of reasons.

I follow Fernandes and Tang (2014) in modeling this uncertainty as a term in the market demand function. Said demand has common uncertainty terms across establishments and an idiosyncratic shock component specific to the establishment. While there will always be uncertainty about idiosyncratic market fit, establishments are able to learn about the permanent (or at least persistent) components of potential market demand for their product. In particular, the functional form of the uncertain demand decomposes demand in a specific market (defined as a product country pair) into a world taste for a specific product category, a country specific taste for American products in general, as well as a country specific taste for a specific product category.

[DETAIL]

Establishments are able learn about these components of demand by observing the export activities of nearby establishments. Specifically, they observe the export status in a specific country-product category, and the average export revenue of the establishment in that category. In the baseline version of the model the export revenue information is observed

[3] Besedeš and Prusa (2011)

without noise, however qualitatively additional noise in the export revenue simply makes the learning process slower. The way in which this uncertainty drives decision making by the establishment is through uncertainty about the expected revenue. Profit π_x from exporting is a function of demand D_{jc} in market jc, defined as the pair of product j and country c and productivity ϕ.

$$\pi_x(D_{jc}, \rho) = D_{jc}\phi^{\sigma-1}$$

where demand is a function of elasticity of substition σ, local wage costs w , idiosyncratic shock Z_{ijc}, price P_{ijc} and market income Y_{jc}

$$D_{ijc} = (\frac{1}{\sigma})^{\sigma}(\frac{\sigma-1}{\sigma * w})^{\sigma-1} Z_{ijc} P^{\sigma}_{ijc} Y_{jc}$$

or, decomposing and in logs,

$$ln(D_{ijc}) = d_{ijc} = \kappa + d_c + d_j + d_{jc} + z_{ijc}$$ [4]

1.2.3 Learning

Establishments are able to infer demand from their neighbors exporting activity. I assume that establishments are able to observe the average of their neighbors exporting revenue and the number of exporting neighbors to a specific market. Establishments then update their prior, normally distributed, beliefs about the demand shock matrix. While

[4]

$$\kappa = ln\left[\left(\frac{1}{\sigma}\right)^{\sigma}\left(\frac{\sigma-1}{\sigma w}\right)^{\sigma-1}\right]$$

updating their beliefs about market demand, establishments decompose demand into product specific, country specific, and product-country specific terms.

$$d_m = d_c + d_j + d_{jc}$$

Relying on the fact that the sum of independent normal random variables is itself normal, and the variance is sum of the variances. Exporters use their initial knowledge of their neighbor's productivity ϕ

Establishments infer past period market demand

$$\bar{d}^{nb}_{m,t-1} = \frac{(\bar{R}_{m,t-1}/n_{m,t-1})}{\phi^{\sigma-1}}$$

Establishments then take this past signal and update their prior for market demand. Since the establishment has three sets of information to use in its learning procedure, it first generates an intermediate belief about market demand based on the signals which match along both the country and product dimension, since that information affects all three components of demand.

$$\bar{d}^{int}_{mt}(n_{m,t-1}, \bar{d}^{nb}_{m,t-1}) = E[\bar{d}^{nb}_{m,t}|n_{m,t-1}, \bar{d}^{nb}_{m,t-1}] = \delta_t \bar{d}^{nb}_{m,t-1} + (1-\delta_t)\,\bar{d}_{jc}$$

$$\bar{d}^{int}_{mt}(n_{m,t-1}, \bar{d}^{nb}_{m,t-1}) = \frac{n_{m,t-1}v_{dm}}{v_{zm} + n_{mt}v_{dm}}\bar{d}^{nb}_{m,t-1} + \left(1 - \frac{n_{m,t-1}v_{dm}}{v_{zm} + n_{mt}v_{dm}}\right)\bar{d}_{jc}$$

where following DeGroot (2004) the aggregated signals provide sufficient statistics to compute the posterior in one step

$$\delta_t(n_{m,t-1}, v_{dm}, v_{zm}) = \frac{n_{m,t-1} v_{dm}}{v_{zm} + n_{mt} v_{dm}} = \left(1 + \frac{1}{n_{m,t-1}} \frac{v_{zm}}{v_{dm}}\right)^{-1}$$

and

$$v_{mt}(n_{m,t-1}, v_{dm}, v_{zm}) = \frac{v_{zm} v_{dm}}{v_{zm} + n_{m,t-1} v_{dm}} = \left(\frac{1}{v_{dm}} + \frac{n_{m,t-1}}{v_{zm}}\right)^{-1}$$

The weight on new signal is increasing in n and decreasing in the variance v_{zm}. The posterior variance of the belief is increasing in the variance of the fundamental and the firm specific demand component, and decreasing in the number of firms. The establishment repeats the same procedure detailed above for each of the remaining components of demand. Establishments update about the product and country specific parameters, not reusing observations. This step is order invariant.

$$d_i^f = E[d_i | n_{i,t-1}, \bar{d}_i^{post}] = \omega_t \bar{d}_{m,t-1} + (1 - \omega_t) d_i^{post}$$

where

$$\omega_t(n_{i,t-1}, v_{di}, v_{d(-i)}) = \frac{n_{i,t-1} v_{di}}{v_{d(-i)} + n_{it} v_{di}} = \left(1 + \frac{1}{n_{m,t-1}} \frac{v_{d(-i)}}{v_{dm}}^{-1}\right)$$

for $i \in \{c, j\}$.

Note that the set of potentially informative firms is neatly partitioned into three groups. The new posterior which enters the firm's decision making is obtained:

$$d_m^{post} = d_c^{post} + d_j^{post} + d_{jc}^{post}$$

$$d_m^{post} = \omega_{c,t}\bar{d}_{c,t-1}^{post} + (1-\omega_{c,t})\left[\frac{d_m^0 - d_c^0}{d_m^0}\right] d_m^{int} + \omega_{j,t}\bar{d}_{j,t-1}^{post} + (1-\omega_{j,t})\left[\frac{d_m^0 - d_j^0}{d_m^0}\right] d_m^{int} + \left[\frac{d_m^0 - d_{jc}^0}{d_{jm}^0}\right] d_m^{int}$$

Simplifying Assumptions

In order to keep the baseline model tractable, establishments only keep updated estimates for the variables which affect their immediate decisions. This limits establishments from attempting to aggregate information about the whole matrix of product and country specific shocks. Theoretically an establishment would want to update all elements of world demand in order to improve its estimates of relevant product components, however this is computationally intensive for both the establishment and the researcher. Recent research about information constraints and costly information acquisition for managers makes this a reasonable assumption.

The baseline version also assumes that all shock components are permanent and immediately fully revealed upon entry. A version with persistent but not temporary skills would operate the same qualitatively, as long as the shock is persistent enough to provide predictive power for future demand.

I also make the assumption that the per period fixed cost of staying in an export market is independent of the spillover variables. This distinction is not crucial to explain any of the empirical findings, but it makes the model predictions about spillover effects more direct.

1.2.4 Export Entry Decision

Establishments enter into exporting if their expected revenue is higher than the initial fixed cost of exporting and then higher than the per period fixed cost for each period going forward.

$$V(entry = 0, \phi, d^0) = \max\left\{0, V(entry = 1, \phi, d^{post}) - f\right\}$$

Since domestic and export profits are additively separably the establishment considers the export decision independently.

$$\pi(\phi) = \pi_d(\phi) + max\{0, E[\pi_x(\phi)]\}$$

The expected profits of the establishment in its export market depends on a number of firm specific fundamentals, the fixed costs, and the inferred demand distribution in the export market.

$$E[\pi_x(\phi)] = \phi^{\sigma`-1}E[D_{im}] = exp(\kappa)\phi^{\sigma-1}exp(\bar{d} + \frac{v_m}{2})$$

Export Productivity Cutoff:

The zero profit condition on entry gives rise to a lower bound on establishment productivity necessary to enter the market.

$$E[\pi(D_m, \rho)] = f_x(S_m)$$

Implying this (posterior) lower bound on productivity

$$\tilde{\phi}(d^{post}, v_{mt}) = \underline{\phi}^{\sigma-1} = \frac{f_x(S_m)}{exp(k) * exp(\bar{d}_m + \frac{v_m}{2})}$$

As shown below, this lower bound on establishment productivity required to justify entry into exporting is decreasing in the posterior belief of the establishment about the potential market demand. A slightly counter-intuitive aspect of the productivity cutoff is that it is decreasing with respect to market variance. This is due to the well known experimenting incentive with log normal shocks. This effect could theoretically lead to the sign on the precision to be ambiguous.

$$\frac{\partial ln\tilde{\phi}}{\partial \bar{d}_{ijc}} = \epsilon_{ijc} = -\delta_t(n_{m,t-1}) = -\left(1 + \frac{v_{zm}}{n_{m,t-1}v_{dm}}\right)^{-1} < 0$$

The model also allows us to derive the property of how responsive the productivity cutoff is to the posterior estimate of market demand, depends on the value of the revenue information and the number of signaling establishments:

$$\frac{\partial |\epsilon_{ijc}|}{\partial n_{m,t-1}} = \frac{v_{zm}}{v_{dm}}\left(n_{m,t-1} + \frac{v_{zm}}{v_{dm}}\right)^{-2} > 0$$

The same inferred demand also pins down a certain initial exporting quantity

$$q_{x,m} = E[R(D_{im}, \phi)]/p(\phi)$$

Survival

The chance that the demand shock $z_m$is large enough that the per period profit is larger than the per period fixed cost to export is isomorphic to the probability of the continuation of the export spell $P_s(\phi, d_m^*)$

$$\frac{\partial P_s(\phi, d_m^*)}{\partial d_m^*} = \frac{1}{\sqrt{v_{zm}}}\phi\left(\frac{1}{\sqrt{v_{zm}}}\left(\ln\left(\frac{f_x}{\phi^{\sigma-1}}\right) - d_m^*\right)\right) > 0$$

Model Implications

The model provides a number of implications which are key to understanding the empirical approach. The relationship between the sign of the revenue signal and the precision of that signal directly motivates the empirical analysis in the following section. The clear partition of the full information set generated by the two different dimensions of exporting, as well as their intersection, adds to the explanatory power of the model when examining the dynamic decisions which operate along one specific dimension of the market.

Notably, the model contains all of the elements which other models highlight in their effect on export participation. The probability of entry is increasing in the establishment's own productivity. It is also increasing with lower fixed costs of exporting. The empirical approach will control for all of these factors, in order to focus on the remaining lever being pulled in the model, the learning dimension. As establishments observe higher market specific average export revenue from neighboring establishments the probability of entering into that market increases.

The model does not distinguish between information coming from different component shocks of the total market demand. The establishment decision is ultimately made using the final posterior estimate of market demand, regardless of which dimension informed the largest portion of that update. This means that the model is ex ante ambivalent about the predicated relative magnitudes of the coefficients on the learning coefficients between information that matches only along one dimension, either only same-country or only same-product.

However, other establishments in the same product same country category provide the most amount of information towards the posterior, as the establishment should weight their information more highly. The model does not take into account any potential discouraging effect from the existence of the incumbent, both domestically and in the foreign market, so empirically it is entirely possible to see a different ranking, as I do, but the model predicts that same country-same product establishment spillovers should be largest.

This also makes intuitive sense. The idiosyncratic shock is the only source of pure noise in the information coming from establishments matching along both dimensions. On the other hand, the establishment perceives the two demand shocks which it does not care about as additional noise. An establishment wanting to learn about a country specific demand shock has to filter out the irrelevant product shock, as well as the irrelevant country-product shock. Mechanically, the variance of the idiosyncratic shock alone is less than the variance of the idiosyncratic shock plus the demand shock associated with the irrelevant market dimensions, i.e. $v_z < v_{-i}$.

The model predicted sign on the precision coefficient is ambiguous. More precise information about a specific market only motivates an establishment to enter there if the information is positive. Similarly, negative information makes it more likely that the establishment would not enter the market. When combined with the experimenting incentive of posterior demand variance embedded in the entry condition, the sign on the precision coefficient itself is not of economic interest.

The model does deliver an unambiguous prediction about how the effect of more precise demand information on the entry probability interacts with the demand information itself. As the establishment's perceived variance of the total market demand, v_{dm} decreases, the establishment places additional weight on the signal. This means, at a given level of the signal, a more precise signal leads to a stronger response of the estimated entry barrier, and therefore a higher entry probability. This effect will be captured by the interaction term on the accuracy and demand value terms in the empirical section. Given that the perceived variance of the total market demand is decreasing in the number of firms, this motivates the importance of including a measure of the number of signals from which the estimate of market demand is derived.

Because the simple model assumes immediate and full revelation of the information upon entry, any establishment that entered into a market would decide in the next period whether to exit or continue indefinitely. In the data we observe many establishments that engage in intermittent exporting, meaning that they exit a market and then re-enter the same market at a later date. Therefore, to test for exit matching the mechanism described in the model, the empirical section focuses on final exit from a specific market, meaning no observed re-entry into that market at any time in the future.

The exit rate of establishments in the model is decreasing in the value signal and the accuracy of the information. More positive news about market demand in a specific market makes it less likely that the fixed cost outweighs the revenue. Furthermore, more precise estimates of the demand shock process reduce the likelihood that the realized shock is negative and large enough to induce exit for the establishment. Even with perfect information this model would have some exit by firms who would profit in expectation but receive a negative idiosyncratic shock. However, more precise information reduces the likelihood that the gap between the establishment's belief about market demand and realized market demand contributes to the exit probability.

The model also explains why establishments might switch their export destination markets throughout their life-cycle. Unlike in a full information model, the establishment's information set continues to change as it observes its neighbors export activity. This can lead an establishment to want to change the set of export markets which it services. This learning process about new markets is conceptually identical to learning about markets as a non exporter. The model predictions about the signs of the coefficients on the value signal and the precision thereof are the same as for entry into a new market.

The level of detail in the definition of the market does lead to some additional predictions. By partitioning the information set into three distinct sets, the establishment decision to switch to a specific market should only be affected by information matching that destination market. For example, an establishment which exports the same product but switches its destination country should not be affected in that decision by product specific information, as the product stays unchanged. Therefore, only same country same product and same country information should matter to this establishment.

This same logic also allows us to rule out the effect of common supply shocks as driving the entry behavior. If a positive supply shock to the county was driving overall export behavior, then we would expect to see a relationship between information which should be irrelevant, which we do not observe.

1.3 Data

In order to more generally answer the question about spillover effects in the U.S., however, it is necessary to assign an export transaction to a specific establishment, which gives it a location. The most detailed source of US microdata on exports, the Longitudinal Foreign Trade Transactions Database (LFTTD), records the universe of goods export transactions at the firm level. For each US export transaction, the LFTTD contains information about the value and quantity traded, the exact date of the transaction, the product code at the HS-10 level, and the destination country. The dataset reports whether the transaction happens at arm's length or between related parties.

This analysis focuses on arm's length exports only, as within firm transactions are unlikely to be affected by learning mechanisms. The Longitudinal Business Database (LBD), with information about the firm and each of its establishments, provides the link between the export transaction and the set of potential establishments that made the export. The LBD also contains information about annual employment and payroll. This lack of detail about establishment level trade in the US Census micro-data is why previous work covering the US in this space has focused on single establishment firm level exports originating from U.S. firms.

In order to do this, a link between the universe of firm level trade transactions from the Longitudinal Foreign Trade Transactions Database (LFTTD) and an establishment within that firm as tracked by the Longitudinal Business Database (LBD) is required. The basis of the approach is to use a series of decision trees combining information about three different key aspects of the establishment and firm to correctly allocate the transaction.

The three most important pieces of information used are geographic information, production associated industry (PAI) information, and survey information in the Annual Survey of Manufactures (ASM) as well as the Census of Manufactures (CMF). This is a variable in the LFTTD which identifies the origin of movement of an export transaction. While this origin of movement variable can be potentially different from the origin of production in the case of warehousing, we account for this by checking the PAI of the establishments. In cases where this matches uniquely to an establishment within a firm, this is a first indication of the shipment's production location.

The second critical piece of information is the production associated industry. Using the Census of Manufacturers Products Trailer File and a rarely used mapping between HS product categories and NAICS industry codes produced by the US Census Bureau, we can associate establishments to commonly produces products. When this now provides a unique link between an export shipment and an establishment within a firm, this is another piece of information that can be used to differentiate the origin of the export transaction.

A survey indicator in the ASM/CMF for establishments also provides an export flag. In the case where two establishments might meet the above criteria, but one indicates in the survey that it exports and the other does not, that is an additional piece of information that can be used in the allocation. The combination of all of these pieces of information leads us to be able to allocate most export transactions in the US to the establishment level. For more detail on the data, see Appendix A and the forthcoming Boehm, Pandalai-Nayar, Flaaen, and Schlupp (Boehm et al.) which describes the allocation process and some initial applications. I use this allocation for the years 1993 to 2015.

The importance of the establishment allocation is shown in two different ways below. First, the importance in capturing the dollar value of all exports is shown in the first figure. It shows that single establishment firms account for roughly 20% of the dollar value of exports each year. The allocation captures the remaining 80% with the other allocation methods described earlier. This chart also shows the difficulties inherent with attempting to generate the same set of local information with only single unit establishments. This difficulty is exacerbated by the selection effect given that single establishment firms are very different from the overall sample.

This can be contrasted with the figure summarizing the number of establishments in a firm per export transaction. For arms-length trade, the focus of this paper, 81% of export transactions are accounted for by single establishment firms. This shows how common small export transactions are in the US. However, even here 20% of the export decisions are not accounted for. This really highlights the importance of having establishment level trade for both sides of the export decision, that is both the decision to export and the local information set observed when making that decision.

The geographic unit of analysis is at the FIPS county level. This allows for sufficiently large variation along the product and destination country dimension, while being small enough that firms can plausibly learn about their neighbors. As of today, robustness checks varying the level of geographic aggregation are still waiting on disclosure approval.

Each unit of observation in the data is an establishment-product-country triad. The data-set contains 33 million establishment-product-country observations from 1993-2015.

Figure 1.1: Share of Arms-Length Exports by Method of Allocation

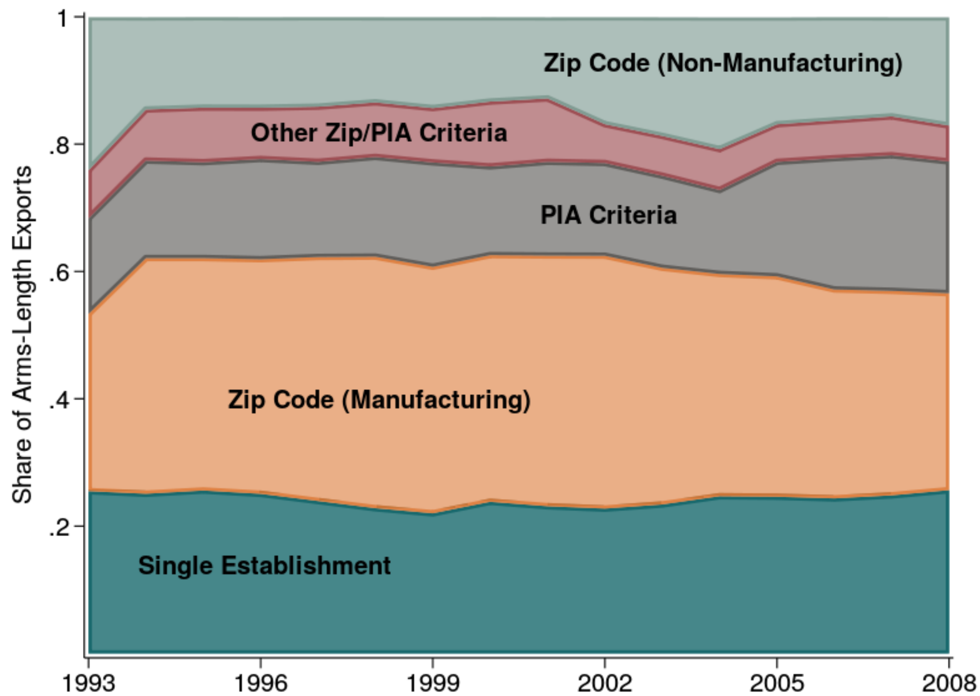


Figure 1.2: Number of Establishments per Export Transaction

Number of Establishments	Share of Transactions
Arms-Length	
1	81 %
2	8 %
3-5	6 %
6-10	2 %
> 10	3 %
Related-Party	
1	61 %
2	16 %
3-5	12 %
6-10	5 %
> 10	5 %

This lets me generate the set of establishments entering into exporting a certain product to a specific destination for the first time, as well as both the firm's and establishments total export history. I also track the set of establishments which exit from exporting a certain product to a certain country for the last time.

1.4 Empirical Strategy

The empirical strategy I employ in this paper is similar in principle to the one designed by Fernandes and Tang (2014) in their study of Chinese exporting firms. The the primary object of interest is the export activity of other establishments in the same county, both in terms of the count of establishments engaged in a specific exporting activity as well as their export volume. Once one expands the definition of a market to consider the difference between product and country, the model cannot quantitatively distinguish between the learning effects along the two channels for a decision like generic entry into exporting. Therefore the empirical section of the paper generates all dimensions of the potential spillovers in order to compare the effects of different ways to construct the spillover variable.

The model indicates that spillover learning effects can come both from the presence of neighboring exporting establishments, as well as information about how those establishments are performing in their export market. Therefore I construct within county spillover variables for both aspects of local exporting activity. Within each county, I consider four mutually exclusive categories of spillover dimensions: all other exporting establishments, all other establishments that export only to the same country, all other establishments that export only the same product, and all other establishments which export both to the same country and export the same product. Additionally, I construct two different versions of each: a simple count of the number of establishments, and the employment weighted share of establishments in each category in the county. This accounts both for the naive counting of total neighboring establishments engaged in exporting activity, but also allows for establishments to pay more attention to the most visible members of its information set.

I also construct four parallel categories of what I call value spillovers, also constructed to consider all four dimension of market fit. This value spillover is defined as the growth rate of the US dollar value of exports to each market, measured as the free on board (FOB) price. The value spillovers take into account the information inherent in the growth rate of establishment export values to specific markets. This expands the information set from which the establishment can extract information about neighboring establishment export demand.

I consider seven outcome variables. The first set of four outcome variables, or establishment outcomes, describe establishment behavior as derived from a simple model of establishment decision making. The first four variables I consider are: first time entry into exporting, final exit from exporting, growth rate of export volume, and labor productivity growth post entry. I study first time entry into exporting because this is the a priori state in which establishments have the least amount of information about their potential market fit. I examine final exit from exporting for a similar reason. Final exit from exporting is the stage at which an establishment has maximum information about its export fit. Finding learning effects in this scenario would indicate the importance of learning about uncertain demand.

The second set of three outcomes is related to more dynamic establishment exporting phenomena. Occasional exporting is a well known property of US exporters which is difficult to account for in canonical models. I investigate whether dynamic learning considerations are a major factor in establishment decisions to occasionally export. I also consider whether dynamic information acquisition guides establishment level product-destination portfolio decisions. Export composition, both at the country and product level varies significantly over a firm's life-cycle. Therefore I apply the same empirical strategy to an establishment's decision to switch its set of destinations while keeping its products the same, an establishment's decision to switch its set of products while keeping its set of destinations the same, and its decision to alter both. I show that reducing uncertainty about product and market fit is an important driver of an establishment's switching decision.

Each of these outcome variables is then used in the estimation of a linear probability model for the binary outcome variables, or an OLS estimation of the continuous variables. Following Bernard and Jensen (2004) linear probability models are an appropriate choice in this context precisely for the ability to include a plant fixed effect, or in my case, a plant-country-product fixed effect. The following equation describes the empirical equation to be estimated for outcome Y of establishment i exporting product j to country k in a given year t.

$$\begin{aligned} Y_{ijct} &= \beta_0 \text{ spillover}_{ijct} + \beta_1 \text{ value spillover}_{ijct} \\ &\quad + \beta_2 \text{ spillover}_{ijct} * \text{value spillover}_{ijct} + \beta_3 \Lambda_{it} + \text{Year}_t + \alpha_{ijc} + \nu_{ijct} \end{aligned} \tag{1.1}$$

The regression specification simultaneously includes a spillover variable, the corresponding value spillover variable, as well as their interaction term. The vector Λ_{it} contains one year lags of establishment employment, county employment, country import demand,

and lagged firm export status in order to control for potential concerns about endogeneity, as discussed in Bernard and Jensen (2004). Following Koenig et al. (2010) I also control for the size of county, as measured by its employment, to capture any potential crowding out effects on transportation infrastructure. The lagged spillover variable also indicates the slow nature of the learning process. Models with full information about the full matrix of demand shocks would not predict establishments to respond with a one period delay.

The other important term to note is the establishment product country fixed effect α_{ijc}. This fixed effect means that the identifying variation exploited here is the time variation in the spillover variable observed by the establishment with respect to the mean level of agglomeration in its region. This conservative specification controls for any remaining confounding effects like supply shocks common to a region, bilateral distance, and the endogenous variation in levels of agglomeration across regions.

1.5 Results

1.5.1 Entry

The columns in the table describe the establishment level response to the spillover variables, as constructed along the dimension in the column title.

For first export entry decisions establishments rely on signals along all three dimensions. Same country same product information is considered the most accurate by the establishments, as indicated by the smallest coefficient on the interaction term. Similarly the country information is perceived as most noisy by the establishments, highlighting the importance of splitting the market definition into different dimensions. It is possible that the

Table 1.1: First Export Entry

	Country	**Country and Product**	**Product**
Average Revenue Signal	0.0130***	0.0146***	0.0174***
	(0.00189)	(0.00142)	(0.00144)
Interaction	0.00213***	0.00146***	0.00199***
	(0.0003)	(0.000163)	(0.000180)
N	31m	31m	31m

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, SE clustered at county level as per Moulton (90)

wide range of other product demand shocks makes it difficult for the establishment to extract the country level demand information without enough observations. The interaction term measures the rate at which additional observations increase the responsiveness of the At the average signal growth rate of 2%, the same country same product coefficient implies a 0.3 percentage point increase and the same product coefficient implies a 0.35 percentage point increase. The sample probability of first market entry is 2%, which implies that moving an establishment from a completely uninformative county to a county with an average positive signal increases its probability of first entry by a significant margin, about 15% of the average entry rate. Given what we know about the importance of fixed costs and establishment productivity in the export decision (both controlled for here), this implies a substantial role for policy in increasing the information available to establishments.

The probability of observing a final export exit is decreasing in the value of the spillover variable, which indicates that establishments are less likely to enter a market incorrectly when they are better informed. This confirms the implication from the model that exit rates are decreasing with respect to both the value and the accuracy of the information.

Table 1.2: Last Export Exit

	Country	Country and Product	Product
Average Revenue Signal	-0.00841***	-0.0111***	-0.0125***
	(0.00146)	(0.00183)	(0.00109)
Interaction	-0.00144***	-0.00111***	-0.00141***
	(0.000214)	(0.000204)	(0.000136)
N	31m	31m	31m

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, SE clustered at county level as per Moulton (90)

Better informed establishments are less likely to experience a negative shock large enough to induce exit. In particular, any decrease in the variance of expected demand increases the survival probability. The exit effect is relatively smaller than the entry effect, comparing a 0.2 percentage point decrease in exit probability to a 2% sample final exit rate.

1.5.2 Dynamic Considerations

Table 1.3: Any Switch

	Country	Country and Product	Product
Average Revenue Signal	0.124***	0.194***	0.166***
	(0.0117)	(0.0196)	(0.00645)
Interaction	0.0225***	0.0209***	0.0217***
	(0.00197)	(0.00212)	(0.000806)
N	31m	31m	31m

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, SE clustered at county level as per Moulton (90)

Establishments also exhibit dynamic behavior inconsistent with a model of full information. Establishments routinely change the composition of their export portfolio, along both the product and country dimension. In this current version a "switch" is defined as changing any part of the total set of the export decision, while a country switch is defined as keeping the set of products exported the same while changing the set of countries, with a product switched defined analogously. This dynamic experimental behavior is further indication that uncertainty about demand and learning from neighbors can be a significant factor in the export decision.

In the dynamic context, establishments respond most strongly to information from other establishments matching along the country and the product dimension. When making any switch, This information is also perceived as the least noisy, showing the lowest coefficient on the interaction term. Increased signal density is most beneficial for signals which match along only one dimension. At the average signal growth rate, the same country same product coefficient implies a 4 percentage point increase relative to the 15% chance of any switch

Table 1.4: Product Switch

	Country and Product	**Product**
Average Revenue Signal	0.0550***	0.0306***
	(0.000902)	(0.00106)
Interaction	0.00574***	0.00308***
	(0.000122)	(0.000119)
N	31m	31m

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, SE clustered at county level as per Moulton (90)

In the case where establishments keep their set of destination countries the same and change the set of products which they export, establishments respond most strongly to signals from other establishments matching along the country and the product dimension.

Increased information density is again most beneficial for signals which match along only the product dimension. At the average signal growth rate, the same country same product coefficient implies a 1.1 percentage point increase relative to the 4% chance of a product switch.

Table 1.5: Country Switch

	Country	**Country and Product**
Average Revenue Signal	0.0116***	0.0145***
	(0.00564)	(0.000902)
Interaction	0.00258***	0.00191***
	(0.000459)	(0.000122)
N	31m	31m

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$, SE clustered at county level as per Moulton (90)

Establishments respond most strongly to signals from other establishments matching along the country and the product dimension. Again I find that increased information density is most beneficial for signals which match along only the country dimension, as predicted by the model. At the average signal growth rate, the same country same product coefficient implies a 0.3 percentage point increase relative to the 2% chance of a country switch. This is a smaller effect than found in the product switch category. This reinforces the relative magnitudes of the learning effects found in the static context.

1.6 Conclusion

The results I find broadly align with the literature finding spillover effects from neighboring exporters. I confirm the existence of local exporter information spillover effects in

US, similar to other developed countries. Koenig et al. (2010) find a that a one standard deviation movement in the number of firms present in a location leads to a 0.55 percentage point increase in entry. In their sample, Fernandes and Tang (2014) find an average of 0.1 percentage point effect relative to a 0.3 percentage point baseline entry probability from the signal but no additional effects from the number of neighbors. These findings are comparable in magnitude to what I find in the US context for same country same product establishments. However, I find notable differences in establishment responsiveness across the different market dimensions of learning. The multidimensional nature of the establishment choice requires a multidimensional approach to disentangle the various shocks and channels.

The learning mechanism appears to be operating qualitatively similarly across the product and market dimensions, in the sense that information which matches only along one dimension is informative, but significantly noisier than information that matches along all channels. Quantitatively, product information appears to have more bite in informing establishment decisions. These findings emphasize the importance of information frictions and imperfect information in all aspects of the export decision. The findings further highlight that establishments are influenced by additional information for not just their entry and exit decisions but also their dynamic lifetime decisions. It also makes the point that the different channels operating on an establishment's information set require detailed consideration in the analysis of export choice.

The findings also raise multiple policy implications for export promotion policies. For one, it clearly highlights that while uncertainty in demand for products is a large influence on establishment decision making, policy interventions to reduce that uncertainty, either directly through provision of information or through the promotion of additional exporters can also be an effective export promotion tool. This is especially interesting given the difficulty of affecting the other important contributors to export decisions, fixed costs and productivity. The lackluster historical success of US export policies is a good indication that more research is needed in this area.

Overall this also points towards the potential benefits of including information frictions and learning models in other contexts within the international trade literature. As data quality improves and the information set available to establishments similarly expands, the potential to investigate other learning channels opens up. Further investigation of the reasons why the learning process differs across dimensions is also warranted.

Chapter 2

The Local-Area Impact of Exporting

2.1 Introduction

[1] One of the most active and influential strands of research in international trade over the past twenty years has focused on the local and regional effects of export and import activity (e.g., Autor et al., 2013). Yet despite the large expansion in the availability and use of micro-level data, relatively little is known about the trade activity of individual locations in the United States. Studies that focus on the US typically use an approach that pairs the regional industrial structure with industry-level data on imports or exports (e.g., Hakobyan and McLaren, 2016). The reason for this indirect strategy is that common datasets contain very limited information on the destination of imports and the origin of exports. More specifically, while information on the port of exportation or importation is available in common datasets, the travels of goods within the US is typically not known. Firm-level data in general does not solve this problem as large, multi-plant firms account for the vast majority of US trade.

In this paper we present a new method for assigning firms' export transactions to individual plants and thereby identify the specific locations most likely associated with the export activity. Beginning with restricted-access, firm-level export transaction data from

[1] This paper is joint work with Christoph Boehm, Aaron Flaaen, and Nitya Pandalai-Nayar

the US Census Bureau, we utilize several additional pieces of information to "allocate" the transaction to the most likely establishment, and hence precise location, from which the transaction originated. This approach is an improvement on plant-level exporting data from the Census of Manufacturers which excludes certain types of trade and provides no destination or product-level detail. It is also an improvement over public-use state-level export data based on "origin-of-movement" indicators, for reasons we describe in detail below. While a number of important complications exist—the firm of export is not always the firm of production, warehousing, intra-firm export consolidation, and the like—the dataset we create provides researchers a more detailed look at the local patterns of exporting than was previously possible.[2]

Using this novel dataset we document a number of new facts on the composition, heterogeneity, and concentration of export activity in the United States. We find that exports are significantly more geographically concentrated than employment or manufacturing sales. At the county level, this pattern holds both nationally and within states. Furthermore, the degree to which these variables differ in their concentration is heterogeneous across states, which highlights that problems may arise when using employment to impute exports to local areas.

Beyond uncovering these new facts, this paper also offers a fresh perspective on the connection between trade activity and local labor markets. Aggregating the micro-data to the county-level, we conduct two exercises that exploit this variation to illustrate the regional heterogeneity in real outcomes arising from exporting during the Great Recession and the

[2]It is important to note that most of these complications will also affect all publicly available export data.

associated trade collapse. Specifically, we find that counties which were more exposed to exogenous negative trade shocks performed worse in terms of employment, pay, and wages during the Great Recession. In our preferred specification, a one million dollar decline in exports is associated with a decline of around 5 jobs in the county. We further show that these effects come not just from the direct employment effects of manufacturing exports, but also because of local spillovers to non-manufacturing industries.

Similar to the majority of the literature studying local and regional labor markets, these estimates are purged of aggregate effects and cannot be used to infer aggregate effects in a straightforward manner. As a first step at estimating broader equilibrium effects, we show that the impact of the export declines do not stop at the county border. We estimate that proximity to more exposed counties results in additional employment losses – an average decline in exports of one thousand dollars per worker in neighboring counties is estimated to lead to an additional 0.8% employment decline in the county of interest. We document similar effects on wages.

The allocation of US firm-level exports to individual establishments offers the potential for documenting and studying other new features of US trade patterns and the effects on the US economy. Broadly, we see three types of analyses which our new dataset makes possible. First, prior to the construction of our new dataset, exports and imports in the US were only measured at the firm-level. However, many other surveys and censuses are conducted at the establishment level. To use trade data together with these datasets from the US Census, researchers typically needed to aggregate establishment-level data to the firm level—which is often imprecise since not all establishments of a given firm are surveyed. Our new dataset enables researches to instead measure exports at the establishment level, and hence such aggregation is no longer necessary. The analysis can instead be performed at the establishment level where researchers can use the available sampling weights. Some examples of questions that might be better addressed with these establishment-level linked data include efforts to map out global value chains and to shed light on the role of non-manufacturing establishments in the trading process. This latter application could also help to inform the scale and growth of so-called factory-less goods producers in the United States (see Bayard et al. (2015) and Bernard and Fort (2015) for a discussion).

Second, our new dataset allows researchers to make more precise statements about industries' trading activity and the effects of trade shocks on industries. Industry classifications are precise at the establishment level but imprecise at the firm level, because large firms tend to be active in multiple industries. Since our new dataset allows us to measure exports at the establishment level, industries' export activity can be measured more precisely. Lastly, our dataset allows us to measure exports by region more precisely than was previously possible. As our applications illustrate, this allows researchers to document regional properties of export activity and to study the effects of shocks on local labor markets.

This paper is related to several strands of the literature. A number of previous papers in the US and other countries have attempted to construct or study sub-national and local trade patterns. For instance, Santamaría et al. (2020) generate a dataset of regional European trade using a new survey, the European Road Freight Transport survey (ERFT), from Eurostat. In the US, the Commodity Flows Survey (CFS) from the US Census has been used to study intra-US trade (see Allen and Arkolakis (2014) and Atalay et al. (2014)). Our dataset supplements these studies by providing data on foreign exports by product and destination, down to the level of individual establishments and counties.

We expect one of the primary applications of our new dataset to be in empirical work studying the relationship of trade and local labor markets. Our paper is naturally related to a long line of influential research studying local labor market effects of trade shocks. Notable papers in this literature include Autor et al. (2013) and many others (China shock), Topalova (2007) (Indian trade liberalization), Kovak (2010) and Dix-Carneiro and Kovak (2017) (Brazil), Hakobyan and McLaren (2016) (NAFTA), and Benguria and Saffie (2019), Benguria and Saffie (2020) (US-China trade war). This line of research typically constructs local exposure to a shock by using national or state-level data mapped to the local level with employment weights, and often finds large and heterogeneous effects across space. Our data opens up avenues for further research in this direction by providing new measures of local exports. Further, our data allows us to distinguish between related party and non-related party trade, which is likely to be useful to study the effect of multinational activity on US local labor markets. For questions in international macro focusing on the transmission of shocks, our dataset has the advantage of containing monthly or even daily export data for local regions.[3]

Finally, our paper is also related to the large literature studying the trade collapse during the Great Recession. This literature explores reasons for the trade collapse, for instance, the role of intermediate input trade or other demand factors (Levchenko et al., 2010; Bems et al., 2013; Bussière et al., 2013) or analyses the role of inventories (Alessandria et al., 2010b,a).[4] To the best of our knowledge, no work has directly assessed the local

[3]See also Caliendo et al. (2019) and Caliendo et al. (2018) for quantitative models studying local labor markets and general equilibrium implications using employment share based approaches for motivating patterns.

[4]The literature on the trade collapse of the Great Recession is vast, including work by Eaton et al. (2016), Ahn et al. (2011), Bown and Crowley (2013), and Feenstra et al. (2014), among many others.

labor market effects of exporting during the trade collapse. In closely related work, House et al. (2019) study the geographic heterogeneity in the effects of exchange rate shocks based on heterogeneity in industry composition and state trade flows, and find empirical evidence consistent with our results for the trade collapse.

2.2 New Data on Local-Area Exports

This section provides details on the new data constructed and used in this paper. We summarize the methodology for the establishment-level dataset of export transactions, discuss the relationship to similar publicly-available data, and document some notable characteristics of local area exports. The existing published data on state-level exports utilizes some select geographical information included in the Customs documentation of an export transaction. While we will also utilize this information, we describe below how, by itself, this indicator could generate misleading results.[5]

2.2.1 Identifying Local-Area Exports

The new data we highlight in this paper build on a number of restricted-use Census Bureau datasets that have become important resources for economists studying firm-level decisions. The Longitudinal Business Database (LBD) is a core piece of the Census microdata as it contains a longitudinally-consistent record of the universe of all business establishments in the US. For our purposes, this dataset provides the industry, employment and payroll of

[5]While customs forms include a destination state and firm for imports, allowing for the construction of publicly available state-level import data, our methodology does not at this stage translate into an allocation of firm-level imports to individual establishments. Some of the steps we describe below that are used in our methodology cannot currently be applied on the import side.

these establishments.[6]

A second underlying core piece of the data that we develop is the firm-level transaction data from the Longitudinal Foreign Trade Transactions Dataset (LFTTD). Constructed from a partnership between US Customs and the Census Bureau, the LFTTD contains the universe of US trade transactions that has been linked to firms via an Employer Identification Number (EIN) as well as other sources. The LFTTD contains the date, value, quantity, and detailed product information (HS10) of a trade transaction along with whether the transaction was conducted between related parties or at arms-length. Finally, we utilize the Census of Manufacturers (CM) "Product Trailer File" to link the products produced by establishments with product-level trade transactions from the LFTTD.

An important point of emphasis is that the LFTTD export data matches an individual trade transaction to a *firm* identified as the US principal party in interest (USPPI); it does not match an export transaction to an individual establishment or location within that firm. Our methodology uses three primary indicators to identify the establishment within that firm that would most likely be associated with a given export transaction. Each of these indicators has advantages as well as limitations, and thus by combining this information together we can arrive at a more accurate distribution of exports across the locations of the firm.

First, the Census of Manufacturer's (CM) Products Trailer file indicates the establishments that produce individual goods in a particular Census year. We use this information to construct, for each exported product, a set of "Production-Associated Industries" (PAIs).

[6]For a detailed description of this dataset, see Jarmin and Miranda (2002).

This product-industry mapping can be used to narrow the potential establishments of a firm that are likely producers of a given exported product (as each establishment of a firm has a unique industry classification). Naturally, some export transactions are associated with wholesale firms with no manufacturing establishments, and hence this method will not apply to all exports. This is an inherent feature of firm-level export data as the firm associated with the export decision may not always be the firm associated with production.

Second, we utilize a variable in the LFTTD that identifies, for each transaction, the geographic "origin of movement": where the shipment began its journey to the port of export. Of course, this origin of movement need not always be the same as the origin of production, as discussed in some detail in Cassey (2009). The three most salient concerns are:

- Consolidation: if a shipment combines with similar commodities from the same USPPI then only the location of consolidation is recorded for that shipment. This is most commonly a problem with agricultural shipments.
- Wholesale/Retail: if the exporter is a wholesaler or retailer, then the location of origin of the wholesaler/retailer is recorded, and not the location of production.
- Warehousing: a shipment could be sent to a storage or warehouse facility and then subsequently exported. In the case where the export process begins at this location, then the origin of movement will be recorded as the warehouse facility, and not the location of production.

Thus, while the origin of movement variables provide useful information for the allocation of exports to plants, these limitations demonstrate the importance of combining it with other information where relevant. Notice that this information is the primary source of information underlying published state-level exports.[7]

Finally, we leverage the establishment export variable as reported in the CM and/or Annual Survey of Manufacturers (ASM) to assist in identifying likely export establishments within a firm. In this last step, where other data are silent, we utilize the share of export activity across a firm's establishments from the CM or ASM to allocate that firm's exports.

Appendix .4.2 provides the full details for how we use each of these sources of information to identify the most likely establishments associated with a firm's export transactions. As shown in Appendix Figure 5, the result is a large fraction of overall exports with a high degree of confidence in the assignment to the correct establishment.

The resulting dataset provides greater granularity and detail of US export transactions than has been previously available to researchers. The wealth of details on individual export transactions—such as product, destination, value, quantity, and arms-length status—are mapped to a precise location of origination for the first time.

2.2.2 Basic Features and State-Level Detail

The only comparison of this data to a publicly-available analogue would be to the Census Bureau's state trade data which is based solely on the origin of movement variable.[8] Compared to a similar industry-product definition – as our allocation methodology removes agricultural and mining products – the disclosed exports based on our methodology account

[7] See `https://usatrade.census.gov/`.

[8] Some researchers study exports at a more granular level using the Commodity Flow Statistics (CFS) data. Unfortunately, this data is only a survey, is infrequently administered, and generally lacks country of destination detail.

for roughly 94 percent of the value identified in the Census public use data product in 2007. The principal reason for the discrepancy between the two numbers is due to trade transactions in the LFTTD that lack a firm identifier, without which there is no establishment-level detail to assign the trade. There are a number of reasons why the researchers that construct the LFTTD are unable to assign all transactions to Census level identifiers that are described in greater detail in Bernard et al. (2009).

Despite the limitations from a conceptual perspective of the publicly available data, the correlation of state-level exports between our measure and the public Census Bureau data product is high at 0.94, which partly reflects the overlap in data sources. Figure 2.1 demonstrates this correlation visually in a scatterplot, showing the (logged) share of state-level exports in the US total between our measure and the publicly available data. Though the logged shares provide a broad view of all 50 states, it masks the few areas where the differences between the publicly available data and our data are significant: Texas (6 percentage points) and Florida (1 percentage point), and Michigan (-3 percentage points). The possibility of over-counting exports at states with ports and under-counting industrial areas is consistent with the drawbacks of the origin of movement-based approach highlighted above, and demonstrates the value-added of combining these other sources of information to provide a more accurate match location.

Table 2.1 shows the top and bottom five largest discrepancies in percentage point terms between the public use data and the data used in this paper. Notable, this list features both small states such as Louisiana, and large states such as California and Texas.

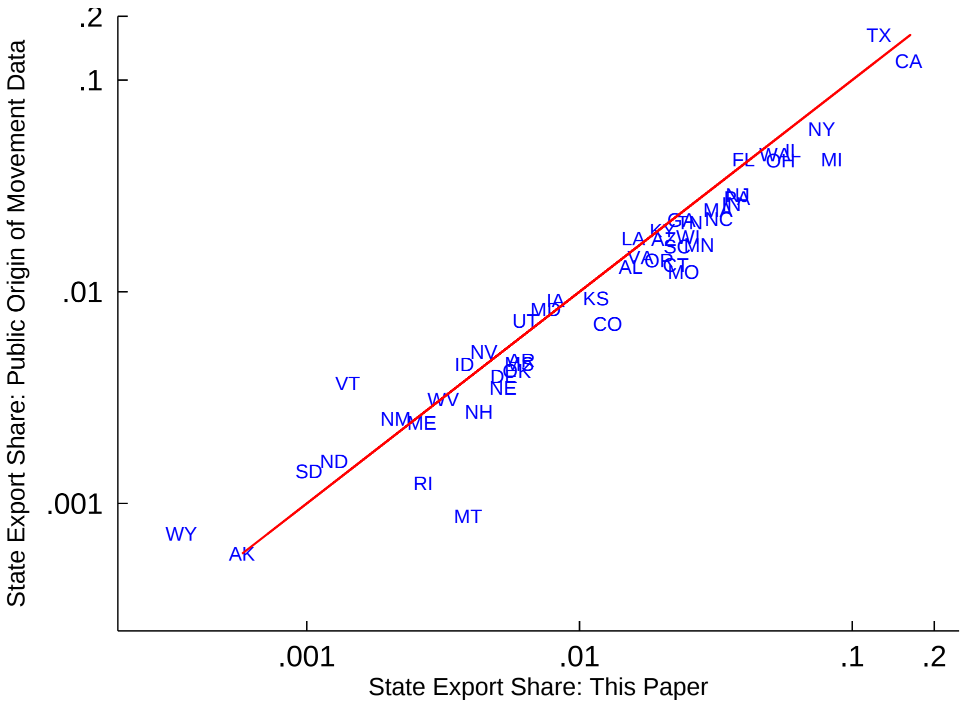
Digging further into the state-level export patterns from our data reveals other impor-

Table 2.1: Top And Bottom 5 Public Use Data Vs Allocated Data

States	Shares (in percent) Public Data	This paper	Percentage point Difference
	Top 5 states: Public Data > Data in this paper		
Texas	16.3	10.3	6.0
Florida	4.2	3.3	0.9
Louisiana	1.8	1.3	0.5
Kentucky	1.9	1.7	0.3
Washington	4.5	4.2	0.3
	Bottom 5 states: Public Data < Data in this paper		
Illinois	4.7	5.2	-0.5
Minnesota	1.7	2.2	-0.6
Missouri	1.2	1.9	-0.7
California	12.3	13.2	-0.9
Michigan	4.2	7.0	-2.8

Sources: Author's calculations using LFTTD, CMF as explained in the text, and Census Bureau State-Level Trade.
Notes: A positive percent difference indicates that the public use data attributes more exports to a specific state than the data used in this paper. Hawaii is excluded for disclosure purposes

Figure 2.1: State-Level Alignment: Our Measure and Census State-Level Exports



Sources: Author's calculations using LFTTD, CMF as explained in the text. Census Bureau State-Level Trade
Notes: This figure provides the share of state-level exports in total exports (in log scale) between our measure (in the x-axis) and the state-level trade database maintained by the US Census Bureau.

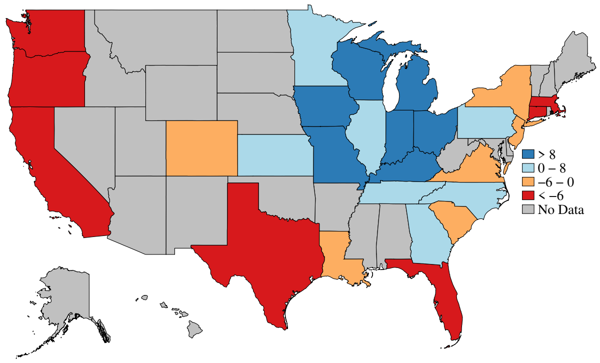
tant features of the geography of US exporting. Beginning with a relatively broad perspective, Figure 2.2 plots the distribution across states of exports to select countries. To draw out the specialization of export destinations across states, the maps report the differential share of exports to a particular destination, relative to the US as a whole. Specifically, for exports of state i to destination k we report:

$$\frac{x_{ik}}{\sum_j x_{ij}} - \frac{X_k^{US}}{\sum_j X_j^{US}}. \tag{2.1}$$

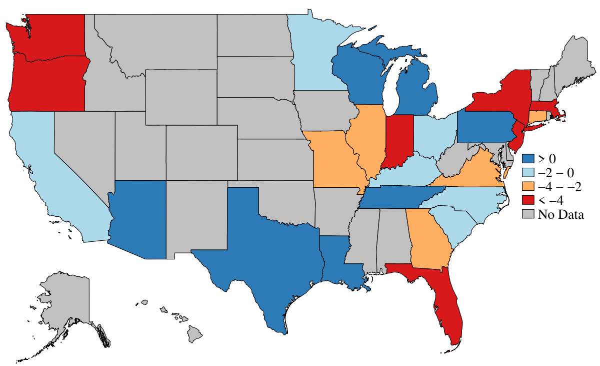
The figure demonstrates the well-known feature that distance matters even apart from country borders, as states close to the Canadian border have a higher-than-average share of

Figure 2.2: State-Level Exports to Select Countries, Percent of Total

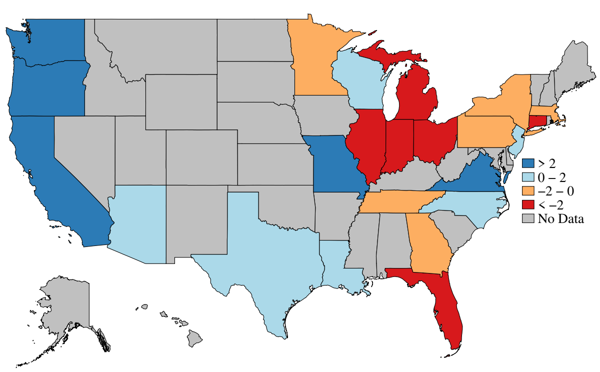
(a) State-Level Exports to Canada



(b) State-Level Exports to Mexico



(c) State-Level Exports to China



Sources: Author's calculations using LFTTD, CMF as explained in the text.
Notes: Each map reflects the differential state-share of exports to each destination, relative to the overall US-level share of exports to that destination. States with grey coloring have data suppressed due to Census disclosure rules. The differential shading reflects the 25th, 50th, and 75th percentiles of each statistic, among those states with disclosed data.

Canadian exports (Figure 2.2a).[9] In a similar fashion, exports to Mexico (Figure 2.2b) tend to be over-represented in states along the southern US border, and exports to China (Figure 2.2c) are higher than average for states along the West Coast. On the other hand, Figure 2.2 shows that the gravity relationship does not explain all important variation in state-level exporting. One important example is the clustering of North American exports in parts of the Midwest such as Michigan, Ohio, and Pennsylvania.[10]

2.3 The Local Concentration of Export Activity

A useful feature of the new data described in this paper is that it can be used to study exporting at a local level. A key new fact revealed by this perspective is the remarkable concentration of export activity in the United States. While it is well-known that economic activity is highly concentrated across space (see Axtell (2001), Gabaix (1999), Gabaix (2011)) and that export activity is concentrated in a small number of firms (see, for example, Bernard et al., 2007), the geographic concentration of exporting has received less attention.

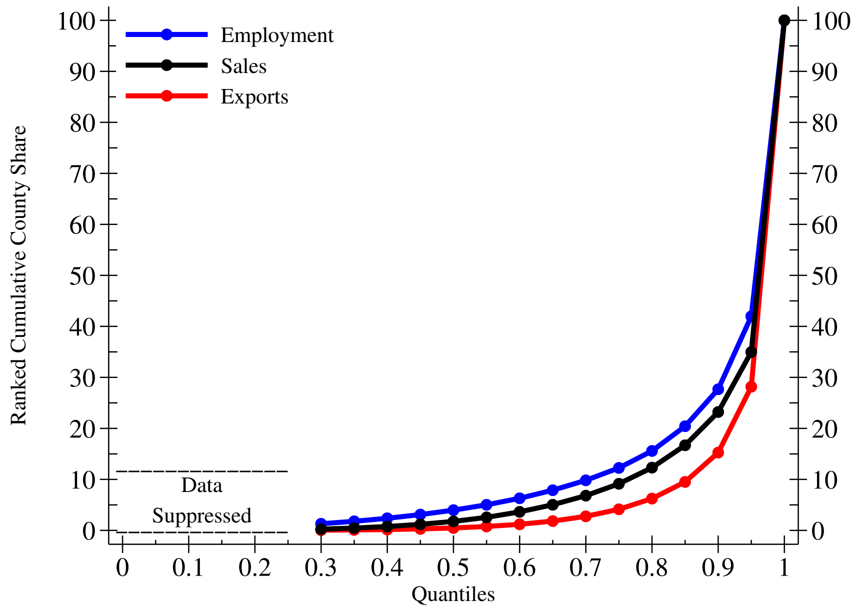
As one way to highlight this concentration, we aggregate our data to the roughly three thousand counties in the United States, rank counties by their share in aggregate exports, and then calculate the cumulative share. For comparison purposes, we conduct a similar exercise for both employment and manufacturing sales, noting that an individual county's rank generally differs for each measure. The resulting figure is similar to a Lorenz curve, but highlights the cumulative share of exports, sales, and employment by the various quantiles

[9]See Crozet and Koenig (2010) for an example with French data, though in their case they use firm-level data and are unable to allocate multi-unit firm exports across *régions* in France.

[10] See Appendix Figure 6 for additional maps corresponding to other export destinations.

of counties in the data. We report detailed quantiles of these ranked shares in Figure 2.3.

Figure 2.3: County-level Concentration: Exports, Manufacturing Sales, Employment



Sources: Author's calculations using LFTTD, CMF as explained in the text.
Notes: These cumulative shares result from ranking US counties according to each variable and then calculating the county-level share relative to the total. Estimates below the third decile are suppressed due to Census disclosure rules.

Figure 2.3 not only confirms the extreme skewness of overall economic activity at this level of geography, but also highlights that export activity is even more concentrated than either employment or sales.[11] While the top 5 percent of counties by employment account for a little over 55 percent of employment, the top 5 percent of counties by export volume account for over 70 percent of total exports.

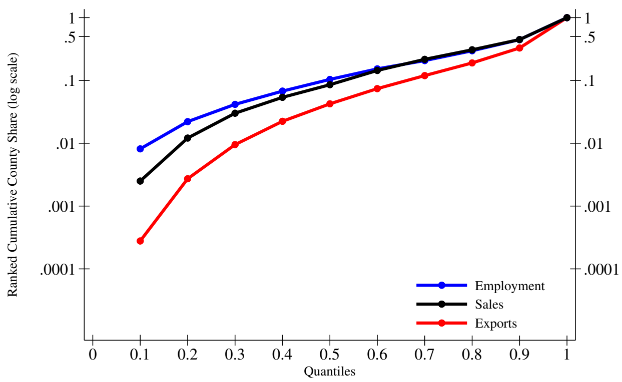
A second and related fact revealed by this data is that much of variation in exporting

[11]The shares of counties below the third decile are so low that disclosure rules prevent us from quantifying their shares.

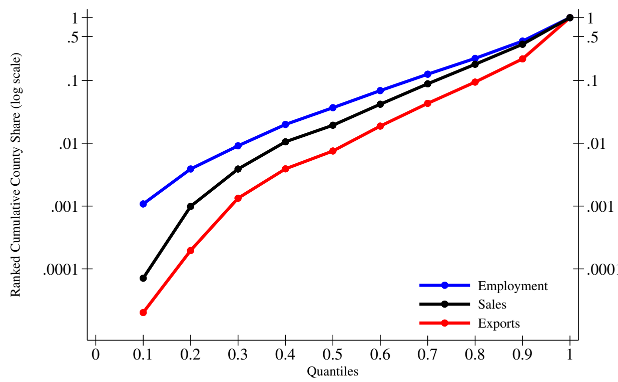
actually occurs *within* states. In Figure 2.4 we conduct a similar exercise to Figure 2.3 but within states rather than across the entire country. Although disclosure limitations prevent a full presentation of all states, we report results for four large states covering different geographic areas of the United States. Given the smaller number of overall counties for a given state, we report these ranked shares by decile, and improve legibility by using a log scale.

Figure 2.4: County-level Concentration by State: Exports, Manufacturing Sales, Employment

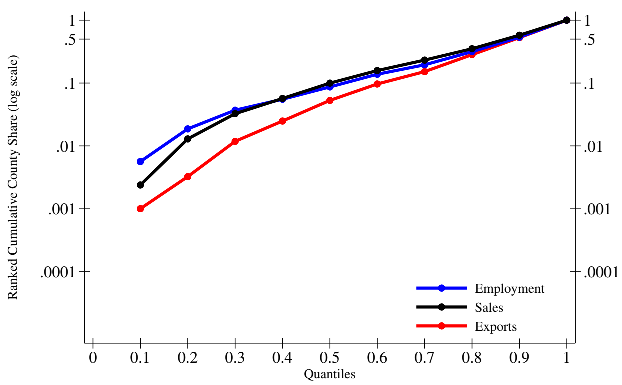
(a) Indiana



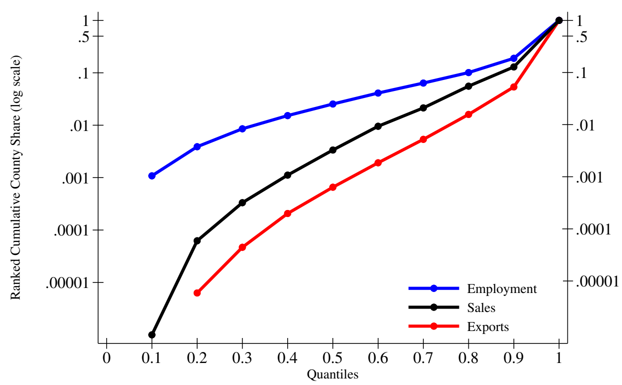
(b) California



(c) Pennsylvania



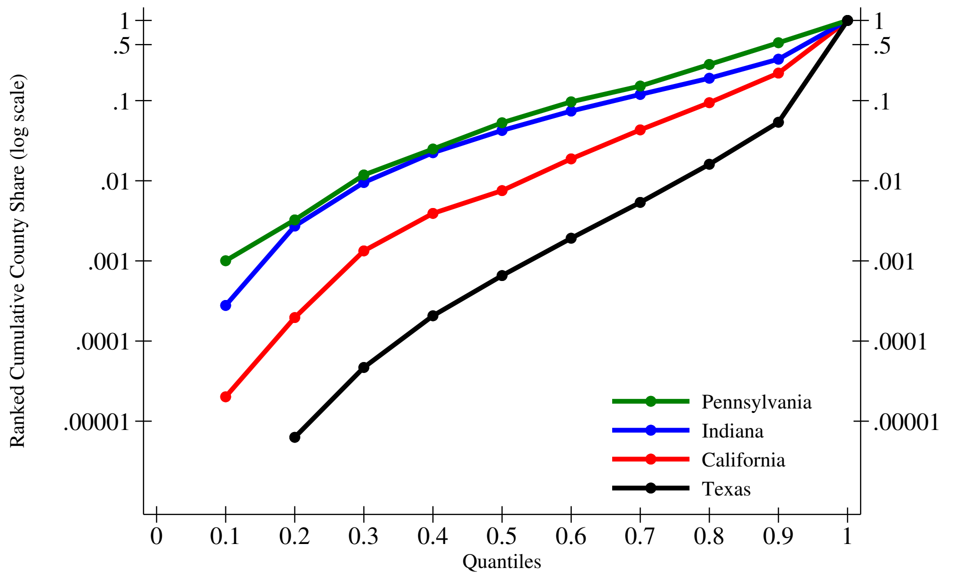
(d) Texas



Sources: Author's calculations using LFTTD, CMF as explained in the text.
Notes: The first exports decile for the state of Texas is suppressed to improve legibility.

Figure 2.4 reveals several notable features. The greater concentration of exports relative to employment or manufacturing sales holds for all states shown, though the degree of this gap differs by state, with the proportions tracking more closely in Pennsylvania than Texas. Second, the degree of export concentration also differs markedly across states. This can be clearly summarized in Figure 2.5 and Table 2.2, which reports the proportion of each

Figure 2.5: County-level Export Concentration: State Comparison



Sources: Author's calculations using LFTTD, CMF as explained in the text.
Notes: The first decile for the state of Texas is suppressed to improve legibility.

variable covered by the top decile of counties in each state.

Table 2.2: Proportion in Top County Decile: Selected States

	Employment	Sales	Exports
Pennsylvania	46%	42%	47%
Indiana	55%	55%	67%
California	58%	62%	78%
Texas	81%	87%	95%

Notes: The table reports the proportion of each measure (employment, manufacturing sales, exports) occupied by the top decile of counties in each state, where the deciles are calculated separately for each measure.

Figure 2.5 demonstrates visually the wide array of concentration of exports across states. Table 2.2 focuses on the variation in the top decile across the illustrative set of states, with as little as 42 percent of activity in the top decile (sales in Pennsylvania), to as much as 95 percent (exports in Texas). In each case, exports are more concentrated than either manufacturing sales or employment.

The stylized facts documented in this section have clearly illustrated the high degree of variation in exports across counties across the United States, and within individual states. This variation does not always align with other indicators of economic activity, with exports typically more highly concentrated than either employment or sales. In the context of evaluating the local labor market effects of trade, this feature is particularly salient in light of the common practice of using county-level employment shares to impute the effects of either imports or export activity to local areas. As our new dataset shows, such a practice could lead to inaccurate interpretations of a local area's export exposure.

2.4 Application to the 2008-2009 Great Trade Collapse

During the Great Recession US real imports and exports both declined by around 20 percent from peak to trough—a much larger drop than overall economic activity. While a sizeable literature has analyzed the factors contributing to this so-called "Great Trade Collapse" (e.g., Levchenko et al., 2010; Alessandria et al., 2010b), data limitations have thus far prevented researchers from studying the effects on US local labor market outcomes. In this section we illustrate the advantages of our new dataset and study how the Great Trade Collapse affected county-level employment, payroll, and wages.

To isolate exogenous variation in foreign demand we employ a version of the World Import Demand (WID) instrument as used, for instance, in Hummels et al. (2014). However, the Great Recession was also a demand shock in the US; hence, to isolate exogenous variation in foreign demand during this period we adapt the WID instrument to only use variation in foreign country-specific declines in imports—and not variation that was common across countries.

We use this variation to estimate whether counties more exposed to foreign demand declines experienced greater employment, payroll or real wage declines between 2007-2009. Our findings in the previous section illustrated that counties differ substantially in their export behavior and hence in their exposure to declines in foreign demand. We therefore expect that the trade collapse during the Great Recession affected countries differentially. Further, we study the direct effects on manufacturing sectors, which account for the large majority of exported goods, and spillovers into non-manufacturing sectors in the exposed counties. We also estimate the effects of geographic spillovers from nearby counties.

There are several reasons why the county-level analysis is preferable over alternative levels of aggregation (e.g., an establishment or state-level analysis). First, a key object of interest to policymakers is the net number of job gained or lost in a given labor market. To the extent that a county roughly aligns with the local labor market, a county-level analysis is informative about this object. In contrast, an establishment-level analysis would measure a gross flow, that is, the number of jobs lost in exporting establishments without taking into account that workers might find jobs elsewhere in the local labor market. While a state-level analysis would also measure the net change in jobs, it would provide very limited variation. Second, a county-level analysis aligns our application with the most common level of aggregation used in previous work that studies the effects of shocks on local labor markets.

We first present our main estimating equation in Section 2.4.1. Threats to identification and the instrument are discussed in Section 2.4.2 and we present the results in Section 2.4.3. Lastly, we consider extensions in Section 2.4.4 and present robustness exercises in Section 2.4.5.

2.4.1 Regression Specification

Let c index counties and t time. We estimate specifications of the form

$$\frac{y_{c,2009} - y_{c,2007}}{\mathrm{emp}_{c,2007}} = \alpha + \beta \frac{\exp_{c,2009} - \exp_{c,2007}}{\mathrm{emp}_{c,2007}} + \text{controls} + \varepsilon_c, \tag{2.2}$$

where $y_{c,t}$ is an outcome of interest, $\exp_{c,t}$ denotes the value of exports, and $\mathrm{emp}_{c,t}$ employment. All variables are aggregated up from establishment-level data in the county. To account for the fact that counties differ in size, we divide both the change in the outcome of interest and the change in exports by the counties' employment in 2007. Notice that this division only affects the interpretation of the constant α; it does not affect the interpretation of the coefficient of interest β.

Our primary outcome variables $y_{c,t}$ are employment and payroll. Employment is measured in numbers of workers and payroll and exports are measured in millions of US dollars. Hence, when the outcome variable of interest is employment, β captures the number of jobs lost due to a drop in exports of one million US dollars. When the outcome variable of interest is payroll, the coefficient β measures the dollars worth of payroll lost after a drop in exports of one dollar. We also estimate equation (2.2) after replacing the left-hand side with the relative change in the average county-level wage $\frac{\text{wage}_{c,2009}-\text{wage}_{c,2007}}{\text{wage}_{c,2007}}$, expressed in percent. In this case we measure exports in thousands of US dollars and hence β measures the percent change in the wage rate attributable to one thousand dollar change in exports per worker.

Specification (2.2) contains a constant and is estimated at the county level. This implies that any aggregate variation over the period from 2007 to 2009 is purged and we estimate the coefficient of interest β entirely from cross-county variation. Note that this coefficient does capture local (general) equilibrium effects such as wage adjustments which may follow a change in foreign demand. As discussed earlier, county-level employment changes following the trade shock reflect *net* changes in the sense that they reflect laid off workers who do not subsequently find jobs elsewhere in the county.

To address the possibility that the outcome variable of interest is trending over time, we control for the lagged change in the dependent variable $\frac{y_{c,2007}-y_{c,2005}}{\text{emp}_{c,2005}}$ in our preferred specifications. We also present specifications in which we control for the 2007 share of employment in exporting establishments, defined as

$$s^{\text{emp,exp}}_{c,2007} = \frac{\sum_{e \in E^{\text{exp}}_c} \text{emp}_{e,2007}}{\sum_{e \in E_c} \text{emp}_{e,2007}}, \tag{2.3}$$

where e indexes establishments, E^{exp}_c denotes the set of exporting establishments in county c, and E_c denotes the set of all establishments in county c. This control aims to address the possibility that counties which differ in their exposure to foreign demand may differ in other unobserved dimensions that could be correlated with developments in their labor market between 2007 and 2009. Lastly, we also present specifications with state fixed effects. In this case the coefficient β is identified entirely from within state variation. We consider additional controls and robustness checks in Section 2.4.5 below.

2.4.2 Identification

The objective of the empirical analysis in this section is to identify the causal effect of a decline in exports—induced by a drop in foreign demand—on counties' local labor markets. Of course, the change in exports on the right-hand side of equation (2.2) and the outcome variable of interest are both endogenously determined and potentially affected by many different shocks. For instance, an adverse local supply shock may reduce both exports and employment and thus bias the estimate of β upwards.

In order for β to capture the causal effect we employ an instrumental variable strategy. In particular, we use a version of the World Import Demand (WID) instrument (Hummels et al., 2014) to isolate variation in foreign demand that is plausibly orthogonal to other shocks that affect county-level exports and employment over this period. Our version of the

WID instrument is constructed as

$$\text{inst}_c^{\text{WID}} = \frac{\text{exp}_{c,2007}}{\text{emp}_{c,2007}} \cdot \sum_{d,p} s_{c,d,p,2007}^{\text{exp}} \left(\Delta \ln \text{imp}_{d,p} - \overline{\Delta \ln \text{imp}_p} \right). \tag{2.4}$$

In this expression $s_{c,d,p,2007}^{\text{exp}} = \frac{\text{exp}_{c,d,p,2007}}{\text{exp}_{c,2007}}$ is county c's export share of product p to destination d and $\Delta \ln \text{imp}_{d,p}$ denotes country d's change in total imports of product p between years 2007 and 2009. Further, $\overline{\Delta \ln \text{imp}_p} = \frac{1}{D} \sum_d \Delta \ln \text{imp}_{d,p}$ denotes the product-specific average of the change in foreign imports, where the average is taken across all possible destinations $d = 1, ..., D$.

At its core, the WID instrument is a weighted sum of changes in foreign imports. It varies by county because different counties are differentially exposed to changes in foreign imports—as captured by the predetermined shares $s_{c,d,p,2007}^{\text{exp}}$. The fraction on the right-hand side of equation (2.4), $\frac{\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$, simply scales the weighted change in foreign imports to account for the fact that some counties export more than others.

A key difference relative to alternative implementations of the WID instrument is that we subtract from the destination and product-specific change in imports, $\Delta \ln \text{imp}_{d,p}$, the product-specific average in imports, $\overline{\Delta \ln \text{imp}_p}$. This adjustment applies and extends the idea of Boehm and Pandalai-Nayar (2020) to purify the WID instrument of undesired components, and addresses the following concern. Since the Great Recession was essentially global in nature, and since downturns affect the demand for products differentially, a version of the WID instrument without this adjustment could be correlated with domestic declines in demand. It is well-known, for instance, that the demand for durable goods declines more during downturns than the demand for nondurable goods. It is also known that a disproportionate share of traded goods is durable. Hence, without purging the instrument of the average decline in demand for products, it likely that the instrument would be correlated with changes in the US demand for the counties' products. This adjustment implies that the instrument as constructed in equation (2.4) uses destination-specific rather than destination- and product-specific variation in foreign demand. In particular, a counties' demand from abroad was disproportionately affected by the collapse in trade if it was unlucky enough to export to destinations whose import demand for its products declined by more than the average global demand for its products during the Great Trade Collapse.

We use annual imports from the United Nations Comtrade database to construct the WID instrument. A product in equation (2.4) is measured at the Harmonized System (HS) 3-digit level. The change in imports, $\Delta \ln \text{imp}_{d,p}$, is constructed as the relative change in imports from all partner countries except for the US.[12]

2.4.3 Baseline estimates

Panel A of Table 2.3 presents the effects on employment. We begin with specification (1), which reports estimates without controls. The estimated coefficient on the change in exports is 4.75, implying that a one million dollar decline in exports to foreign countries led to the loss of 4.75 jobs. This estimate is significant at the one percent level. Little changes when we include the lagged change in employment to account for pre-trends (specification (2), and when we additionally include the 2007 share of employment in exporting establishments (specification (3)). In both cases the coefficient estimate on the change in exports remains stable near 5 and highly significant.

[12]We have alternatively constructed the instrument at the HS 4-digit level in undisclosed results.

In specification (4) we additionally add state fixed effects and therefore identify the coefficient of interest only from within-state variation across counties. In this case the estimate drops to 3.14 but remains significant at the 5 percent level. We report this estimate to demonstrate the robustness of our results but note that care must be taken when interpreting this coefficient. To the extent that declines in foreign demand affect labor market outcomes in nearby counties, the state fixed effect may purge "too much" variation. We provide evidence for the presence of such cross-county spillovers below. For this reason we prefer the estimates from specifications (1) to (3), which suggest that roughly 5 jobs were lost for every one million dollar decline in exports. Lastly, we note that the first-stage F statistics exceed the conventionally applied threshold of 10 by a large margin for all specifications reported in Panel A of Table 2.3. This suggests that the instrument is uniformly strong.

We next turn to the effects on payroll, which are reported in Panel B of Table 2.3. Specification (1) reports the estimate of interest without controls. This estimate is 0.348. In specifications (2) and (3) we successively add the lagged change in payroll and the 2007 share of employment in exporting establishments. The coefficient of interest is stable, approximately equal to 0.35, and highly significant across these specifications. A coefficient of 0.35 implies that a one dollar decline in exports reduces payroll by 35 cents or, put differently, that 35 percent of the decline in exports is passed through to payroll. Again the estimate falls slightly when we include state fixed effects, as reported in specification (4). The same caveats as discussed above apply to this estimate.

Table 2.4 reports the effects on average pay per worker—to which we henceforth refer to as wages. More specifically, the dependent variable is the percent change in wages from

Table 2.3: Baseline estimates: Employment and Payroll

Panel A: Employment

	Dependent variable: $\frac{\text{emp}_{c,2009}-\text{emp}_{c,2007}}{\text{emp}_{c,2007}}$			
	(1)	(2)	(3)	(4)
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	4.75***	4.80***	5.50***	3.14**
	(1.19)	(1.16)	(1.67)	(1.29)
$\frac{\text{emp}_{c,2007}-\text{emp}_{c,2005}}{\text{emp}_{c,2005}}$		-0.029	-0.028	-0.079***
		(0.028)	(0.027)	(0.028)
$s^{\text{emp,exp}}_{c,2007}$			0.000	0.000
			(0.000)	(0.000)
State fixed effects	no	no	no	yes
First stage F	671.0	670.0	474.9	444.2

Panel B: Payroll

	Dependent variable: $\frac{\text{pay}_{c,2009}-\text{pay}_{c,2007}}{\text{emp}_{c,2007}}$			
	(1)	(2)	(3)	(4)
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	0.348***	0.347***	0.372***	0.273***
	(0.047)	(0.046)	(0.059)	(0.048)
$\frac{\text{pay}_{c,2007}-\text{pay}_{c,2005}}{\text{emp}_{c,2005}}$		-0.009	-0.011	-0.086***
		(0.025)	(0.024)	(0.025)
$s^{\text{emp,exp}}_{c,2007}$			0.043	0.025
			(0.071)	(0.043)
State fixed effects	no	no	no	yes
First stage F	671.0	673.8	476.4	448.3

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports. Exports and payroll are measured in millions of dollars and employment in numbers of workers. The employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

2007 to 2009, expressed in percent. Further, for ease of interpretation, exports are now measured in thousands of US dollars. The estimate of interest in specification (1) without controls is 0.744, implying that a decline in exports per worker by 1000 US dollars reduces the county-level wage by 0.744 percent. This estimate is roughly stable when we successively add the lagged change in payroll and the 2007 share of employment in exporting establishments as controls in specifications (2) and (3). When we add state fixed effects in specification (4) the estimate again falls slightly.

Table 2.4: Baseline estimates: Wage

Dependent variable: $\frac{\text{wage}_{c,2009}-\text{wage}_{c,2007}}{\text{wage}_{c,2007}}$ (percent)				
	(1)	(2)	(3)	(4)
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	0.744***	0.789***	0.863***	0.752***
	(0.120)	(0.121)	(0.165)	(0.130)
$\frac{\text{wage}_{c,2007}-\text{wage}_{c,2005}}{\text{wage}_{c,2005}}$		-0.256***	-0.259***	-0.297***
		(0.029)	(0.029)	(0.029)
$s^{\text{emp,exp}}_{c,2007}$			0.000	0.000
			(0.000)	(0.000)
State fixed effects	no	no	no	yes
First stage F	671.0	667.4	468.8	441.9

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports. The dependent variable is the relative change in the wage from 2007 to 2009, expressed in percent. Exports are measured in thousands of dollars and employment in numbers of workers. The employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

2.4.4 Extensions

We next consider two extensions in which we explore how the foreign demand shock during the Great Recession spilled over to other sectors and across county borders.

Manufacturing versus non-manufacturing

We begin with studying the degree to which the foreign demand shock affected manufacturing sectors and non-manufacturing sectors differentially. Since the manufacturing sector produces the large majority of traded goods, 78 percent in 2007 (see Census, 2008, p.1 of FT-900 Supplement), one would expect that a decline in foreign demand affects this sector disproportionately. On the other hand, lost income of manufacturing workers may trigger declines in the demand for locally produced goods and services, thereby leading to contractions in the non-manufacturing sectors through local equilibrium effects. This section aims to estimate the magnitudes of these effects.

To do so we consider the same outcome variables as before—i.e. employment, payroll, and wages—but construct them separately for manufacturing and non-manufacturing industries within each county. We identify manufacturing plants based on their 2-digit NAICS classifications (31, 32, and 33) and refer to aggregates over the manufacturing industries as the manufacturing *sector* and to aggregates of non-manufacturing industries as the non-manufacturing sector.

Table 2.5 presents the results. We report coefficient estimates of our preferred specifications, which include the lagged dependent variable (from 2005 to 2007) as well as the 2007 share of employment in exporting establishments as controls, but do not include state fixed effects. Specifications (1) and (2) show that employment was impacted by similar magnitudes in the manufacturing and non-manufacturing sectors. More precisely, the estimates suggest that a decline of exports of one million US dollars reduced manufacturing employment by 2.72 jobs while it reduced non-manufacturing employment by 3.38 jobs. Both estimates are significantly different from zero at conventional levels, but the effect for the non-manufacturing sector is estimated with lower precision. While the larger magnitude of the effect on employment for non-manufacturing may seem at first surprising, recall that the level of manufacturing employment is considerably lower than non-manufacturing employment in most counties.

Specifications (3) and (4) show the effects on payroll. The estimates suggest that for every dollar decline in exports payroll fell by 24 cents in manufacturing sectors and by 16 cents in non-manufacturing sectors. Since payroll fell by more in the manufacturing sector while employment fell by less, the difference must be accounted for by wage changes. Specifications (5) and (6) confirm this intuition. In particular, specification (5) suggests that a decline in exports of one thousand dollars per worker led manufacturing wages to fall by 1.54 percent. In contrast, this decline in foreign demand had only small and insignificant effects in the non-manufacturing sector.

Spillovers to neighboring counties

We next turn to the question of whether a counties' labor market outcomes during the Great Recession were affected by adverse foreign demand shocks for goods produced in neighboring counties. There are multiple reasons why one might expect this to be the case. For instance, if workers commute across county lines a job loss in an adjacent county may affect income in the county of interest. Alternatively, it is possible that counties' economic

Table 2.5: Effects on Manufacturing and non-Manufacturing Sectors

Dependent variable	Employment		Payroll		Wages	
	Mfg. (1)	Non-Mfg. (2)	Mfg. (3)	Non-Mfg. (4)	Mfg. (5)	Non-Mfg. (6)
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	2.72***	3.38**	0.24***	0.16***	1.54***	0.29
	(0.74)	(1.42)	(0.05)	(0.06)	(0.42)	(0.20)
lagged dependent variable	0.06**	-0.10**	-0.10***	-0.68***	-0.28***	-0.12***
	(0.03)	(0.04)	(0.04)	(0.22)	(0.03)	(0.05)
$s_{c,2007}^{\text{emp,exp}}$	0.00	0.00	0.04	0.02	0.00	0.00
	(0.00)	(0.00)	(0.07)	(0.07)	(0.00)	(0.00)
State fixed effects	no	no	no	no	no	no
First stage F	474.8	475.2	482.7	474.4	475.6	474.0

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports. Abbreviations: Mfg.—Manufacturing; Non-Mfg.—Non-manufacturing. In specifications (1) and (2) the dependent variable is $(\text{emp}^s_{c,2009} - \text{emp}^s_{c,2007})/\text{emp}_{c,2007}$, where $s \in \{\text{Mfg.}, \text{Non-Mfg.}\}$ indexes the sector. In specifications (3) and (4) the dependent variable is $(\text{pay}^s_{c,2009} - \text{pay}^s_{c,2007})/\text{emp}_{c,2007}$. In specifications (1) through (4) exports and payroll are measured in millions of dollars and employment in numbers of workers. In specifications (5) and (6) the dependent variable is $(\text{wage}^s_{c,2009} - \text{wage}^s_{c,2007})/\text{wage}^s_{c,2007}$, expressed in percent, and exports are measured in thousands of dollars. In all specifications the employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. *Lagged dependent variable* refers to the dependent variable from 2005 to 2007. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

activity is coupled through various equilibrium effects. In this section, we explore these connections between counties.

To do so we first construct a measure of the change in exports in nearby counties. Specifically, we proceed as follows:

1. For each county c, identify the closest 10 counties. Then drop those counties that are more than 100 miles away.[13] Index the remaining counties by $n_c = 1, ..., N_c$, where $N_c \leq 10$.

2. Construct the weights
$$w_{n_c,2007} = \frac{\text{emp}_{n_c,2007}}{\sum_{n_c=1}^{N_c} \text{emp}_{n_c,2007}}. \tag{2.5}$$

3. Construct the weighted average of changes in exports relative to employment,
$$\text{neighbor shock}_c = \sum_{n_c=1}^{N_c} w_{n_c,2007} \frac{\text{exp}_{n_c,2009} - \text{exp}_{n_c,2007}}{\text{emp}_{n_c,2007}}. \tag{2.6}$$

We then estimate specification (2.2), where we now include the shock to neighboring counties (2.6) on the right-hand side.

Of course, the change in exports of neighboring counties may be endogenous, just as the change in exports of the county of interest. We therefore construct an instrument for exports from these counties,
$$\text{inst}_c^{\text{neighbor}} = \sum_{n_c=1}^{N_c} w_{n_c,2007} \cdot \text{inst}_{n_c}^{\text{WID}}, \tag{2.7}$$

[13] To measure distances we use great-circle distances calculated with the Haversine formula based on internal points in the geographic area. For details, see the NBER county distance database, which is available here.

where $\text{inst}_{n_c}^{\text{WID}}$ is the WID instrument given by equation (2.4).

Table 2.6 reports the results, beginning with the effects on employment in specifications (1) and (2). The estimates suggest that a change of exports in neighboring counties by one thousand dollars per worker changes employment in the county of interest by 0.8 percent (since exports are measured in millions, the coefficient must be divided by 1000). Notice that the coefficient estimate on exports from the own county is largely unaffected by the inclusion of the shock to neighboring counties.

Specifications (3) and (4) show the effects on payroll. The estimates imply that an average decline of exports by one dollar per worker in neighboring counties led to a decline of almost 40 cents per worker in the county of interest—a surprisingly large effect. Lastly, specifications (5) and (6) show the effects on wages. A decline in exports by one thousand dollars per worker in neighboring counties led to a a decline of wages between 0.540 and 0.725 percent.

Since we now have multiple endogenous regressors and multiple instruments, we report the standard F statistic, as well as the Cragg-Donald F statistic (Cragg and Donald, 1993) and Sanderson and Windmeijer conditional F statistics (Sanderson and Windmeijer, 2016). All of these statistics are large for all specifications relative to conventionally used critical values (for critical values for the Cragg-Donald F statistic, see Stock and Yogo, 2005), implying that the instruments are individually and jointly strong.

Aggregation As is common in most work studying local labor markets, our analysis is purged of aggregate effects. While estimating spillovers from neighboring counties takes a

Table 2.6: The effects of shocks to neighboring counties

Dependent variable	Employment		Payroll		Wages	
	(1)	(2)	(3)	(4)	(5)	(6)
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	4.18***	5.09***	0.320***	0.351***	0.705***	0.828***
	(1.18)	(1.30)	(0.046)	(0.05)	(0.12)	(0.13)
neighbor shock$_c$	7.93***	8.13***	0.384***	0.391***	0.540***	0.725***
	(2.05)	(2.08)	(0.062)	(0.064)	(0.179)	(0.183)
lagged dependent variable		-0.04		-0.025		-0.267***
		(0.03)		(0.025)		(0.029)
$s^{\text{emp,exp}}_{c,2007}$		0.00		0.06**		0.00
		(0.00)		(0.03)		(0.00)
State Fixed Effects	no	no	no	no	no	no
First stage F: own county	210.4	209.2	210.4	208.4	210.4	208.9
First stage F: neighbor	71.0	62.1	71.0	62.0	71.0	61.2
Cragg-Donald F	328.8	237.1	328.8	238	328.8	234.4
First stage SW F: own county	424.6	405.1	424.6	421.3	424.6	421.9
First stage SW F: neighbor	131.9	107.9	131.9	120.2	131.9	119.6

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports and the instrument (2.7) to address the endogeneity of the change of exports in neighboring counties. In specifications (1) and (2) the dependent variable is $(\text{emp}_{c,2009} - \text{emp}_{c,2007})/\text{emp}_{c,2007}$. In specifications (3) and (4) the dependent variable is $(\text{pay}_{c,2009} - \text{pay}_{c,2007})/\text{emp}_{c,2007}$. In specifications (1) through (4) exports and payroll are measured in millions of dollars and employment in numbers of workers. In specifications (5) and (6) the dependent variable is $(\text{wage}_{c,2009} - \text{wage}_{c,2007})/\text{wage}_{c,2007}$, expressed in percent, and exports are measured in thousands of dollars. In all specifications the employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. *Lagged dependent variable* refers to the dependent variable from 2005 to 2007. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level. For details on the Cragg-Donald F statistic, see Cragg and Donald (1993) and Stock and Yogo (2005). SW F is the Sanderson and Windmeijer (2016) conditional F-statistic.

first step at exploring various equilibrium effects, we do not further aggregate our estimates to construct an implied aggregate employment decline due to the trade collapse, as our estimates are not fully informative on this question.

Discussion The analysis this far helps illustrate the benefits of the new data in this paper. The benefits of these data are not limited to the analysis of this shock, but could be used to study the effects of other trade shocks in the past three decades. Further, these data open up the possibility of linking the literature on local labor markets with many important questions in international economics that have largely been tackled at an aggregate level. For instance, these data would permit an analysis of the role of multinational exports or of export links with specific destinations—such as China—on local employment.

2.4.5 Robustness

Due to Census disclosure restrictions, we only present results for a set of primary specifications. However, we have run a number of other robustness checks on these results which we briefly summarize here. First, the results reported above use the smallest common sample of counties for disclosure purposes, but results are stable in sign and significance when using all counties available for each specification. The reported results are also robust in sign and significance to using county level per-worker wage controls to proxy for average productivity within the county. Finally, the results are robust in sign and significance to a Borusyak et al.-style (ming) restatement of the shocks and industry level outcomes.

2.5 Concluding Remarks

This paper presents a novel methodology for allocating export transactions across a firm's establishments. Applying this methodology to microdata from the US Census Bureau results in a new dataset that offers researchers greater detail about the origin of shipments to complement the high-level of detail that transaction-level export data allows. In this paper, as a first pass, we exploit the increased geographic breakdown to document a number of new stylized facts. We show that export activity is more geographically concentrated than either employment or sales. We further document that the differential degree of concentration amongst these variables varies substantially across US states. These features of export concentration pose challenges to the traditional method that assigns trade exposure to local areas based on industry-level employment shares.

The new dataset will have a diverse array of research applications. When applied to the collapse in trade during the Great Recession in the US, we show that the exposure to declines in foreign demand exacerbated declines in employment in local labor markets, both in the county where the export activity was located, as well as in neighboring counties which were indirectly exposed to this trade shock. Because these results are purged of aggregate effects, we leave for further research a full accounting of the relationship between changes in foreign demand and job losses during the Great Recession. We hope this tighter geographic link between trade activity and other economic variables will be useful for future researchers, and in turn, to policymakers eager to understand the impact of trade on the broader economy.

Chapter 3

US exporters between 1993-2015

3.1 Introduction

The evolution of firms throughout their life cycle has been an interesting subject of study ever since high quality longitudinal data has been available at the firm level.[1] [2] This is particularly true for the subset of firms which engage in international exports. While productivity is widely known to be the major determinant of export entry, the dynamics of exporters across their life cycle is nonetheless a very active and interesting field of study. This large but still growing literature dynamically studying exporters, includes e.g. Gumpert et al. (2020) studying multinational exporters in France, Norway, and Germany. Their work is related to line of literature such as Ruhl and Willis (2017) and many others.

However, in particular, the interaction of the timing between firm creation, firm entry into exporting, and firm life cycle performance for firms in the US remains relatively understudied. Many of the existing papers, e.g. Alessandria and Choi (2014) focus on the model implied cost-benefit trade-off structure of exporting, without focusing as much on when in the firm's lifecycle those decisions are made. Using access to confidential US Census

[1]This paper is in progress pending a completed Census disclosure review

[2] Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. DRB approval number: CBDRB-FY20-173

micro-data providing the full panel of US exporters I construct a dataset which tracks both the overall sample of US exporters from 1993-2015 as well as the panel of US firms for which I observe both firm creation and firm entry into exporting with certainty. I begin this paper by documenting a novel set of facts about exporter performance over the life cycle for all exporters and for the fully tracked firms. I then focus on specifications which begin to answer the question: How does a firm's timing in export decisions affect its life time performance?

Deng et al. (2017) have a very intriguing finding from the strategic management literature showing that the effects of early export entry on firms can be quite polarized, following up on Deng et al. (2014). The authors find negative effects from early exporting related to a lack export competition or the inability to attract foreign funding. However, in highly competitive situations where firms export early, those firms tend to be selected from more strategically nimble firms which perform better. This feedback mechanism between export timing and firm creation presents an interesting hypothesis to test in US firms. I find that US firms perform better in a lifetime sense if they began exporting the year they were born as opposed to later. I also find that US firms which engage in multiple types of international trade (both related party and non related party trade) perform better than US firms which only engage in arms length exports.

Additionally, I emphasize the period before firm exit as well as the estimated probability of exit. This aspect of exporter life cycles has traditionally been even less studied than life cycle dynamics in general. Impullitti et al. (2013) write down an interesting model of export entry and export exit for firms in the US. They also find a role for the way in which firms were induced to enter into exporting. A reduction in overhead costs as opposed to a decrease in sunk costs matters for a firm's persistence in exporting. Muûls (2008) argues that credit constraints as well as productivity matter for exporting. It is possible that firms which are able to begin exporting immediately are selected from a less credit constrained population than other similarly productive firms for some reason, leading to better lifetime access to credit. Ilmakunnas and Nurmi (2010) makes a similar point about capital intensity and its relationship to early exporting, noting that more capital intensive firms enter earlier, maybe due to easier access to funding. This paper does not hypothesize a specific mechanism, but focuses on documenting the effects of differential export timing along a firm's life-cycle in order to inform further research into the potential mechanisms by which these observed effects operate.

3.2 Data

From the US Census, this paper used the Longitudinal Foreign Trade Transactions Database (LFTTD), which includes the universe of transaction-level trade in goods from 1993 to 2015. For each US export transaction, the dataset contains information about the value and quantity traded, the exact date the transaction took place, the product code at the HS-10 level, and the destination country. The dataset reports whether the transaction happens at arm's length or between related parties (intrafirm trade) but does not include additional information about the firm's ownership status. In particular, the LFTTD does not identify US-owned or foreign-owned multinationals separately, and has a very low threshold for defining trade as between related parties. In addition, from the Longitudinal Business Database (LBD, covering the universe of US private nonfarm establishments), I obtain information about annual firm employment and payroll, and the age of the firm, which is crucial to perform the life-cycle analysis.

For comparability with the previous literature, I maintain the entire LFTTD sample to construct the cross sectional facts (5,673,000 observations, around 192,000 firms per year), but I restrict the sample to firms which survive for at least five years in the export market to construct the life-cycle facts (59,000 observations). I also generate a balanced panel of firms for which I observe firm creation and the first five years of exporting (8,500 observations and 1,700 firms).[3] I only include manufacturing plants in the analysis. In the US Census data, employment is the most accurate measure of firm size from a time-series perspective. Firm sales are not available for all plants every year, but only in Census years. I therefore present results related to size for firms in the Census data using employment as the size measure. Where feasible, I have conducted robustness exercises using the quinquennial measure of firm sales from the Census of Manufacturers. The results remain qualitatively unchanged.[4] We also control for partial first year effects, see Bernard et al. (2017)

3.3 Static Facts

Table 3.1 shows some broad summary statistics about the two samples used. Among the 5,673,000 observations contained in the Census sample (192,000 firms on average per

[3]The LFTTD begins in 1993. For this reason, I define *export entrants* as firms which do not exporting in 1993 but export at a later year. In this way, I can guarantee that we are observing the first five years of exporting.

[4]More precisely, plant-level sales are available for all manufacturing plants in Census years (ending in '2' or '7'), and for a subset of plants in interim years. In the robustness exercises, I scale up the sales of plants to a firm-level measure using sales by plants that are reported in interim years and using scaling factors based on the share of employment in plants for which sales are reported in interim years.

Table 3.1: Sample Summary Statistics

	All US-based firms	Balanced panel of US-based firms
No. of observations	5,673,000	981,000
with positive exports	2,481,000	291,000
	43.7%	29.7%

Sources: Author's calculations using US Census LFTTD/LBD as explained in the text.
Notes: This table contains firm-year counts of US firms from the LBD, as well as the firm-years during which those firms report positive exports through the LFTTD. The balanced sample refers to the subsection of firms from 1994 to 2014 for which we observe firm creation and potential export entry.

year), 43.7% of firm-year observations export at some point in their life. [5] The disclosure currently pending disclosure review board approval will disclose summary statistics from all samples about various firm characteristics.

Table 3.2: Outcome Variable Summary Statistics

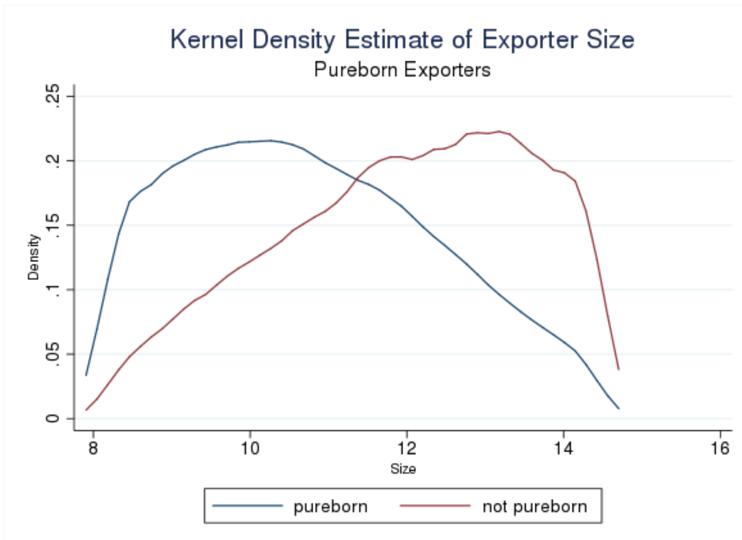
Sources: Author's calculations using US Census LFTTD/LBD/CMF as explained in the text.
Notes: Placeholder table for pending disclosure of summary statistics of all firm level variables

Being able to accurately classify exporting firm's type is important along multiple dimensions. Pure born exporters, which are firms born exporting, and only ever engaging in non-related party exports their whole life, are much smaller than other firms. The following figure shows that knowing whether or not a firm was born as a firm purely focused on

[5]The fraction of firms in the Census reporting exports is higher than say Bernard and Jensen (1999) because I condition on firms that survive a minimum of five years, which excludes many small, typically non-exporting firms.

exporting will have a large implication about its size at birth, and over its life cycle. This type of often unobserved characteristic could, for example, confound analyses about firm size if there is a compositional shift to pure born exporters underlying the sample.

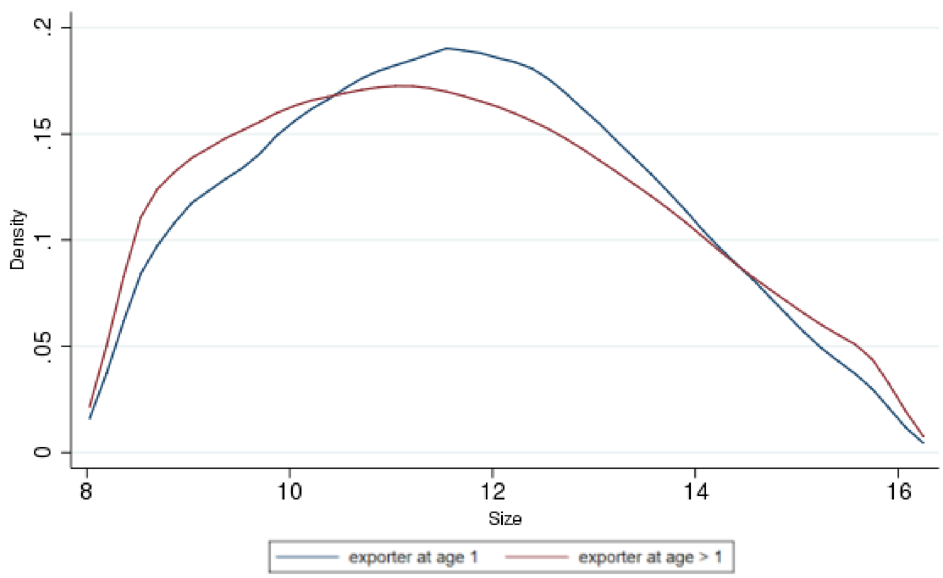
Figure 3.1: Firm Size: Exporter Type Balanced Sample



Sources: Author's calculations using LFTTD/LBD as explained in the text.
Notes: This chart shows the kernel density chart of firm size, defined by firm employment from the balanced sample. The balanced sample refers to the subsection of firms from 1994 to 2014 for which we observe firm creation and potential export entry. Pureborn exporters are firms which are born exporting at arms length and never export to a related party

This size difference across types is also not trivial to predict. In this figure 3.1, showing the size difference of firms which export from the first year they're born compared to firms which export at ages later than the first year, shows that firms which export from the very beginning are larger, on average. This finding is not surprising in the well known context of exporting firms tending to be larger and more productive. However, it also highlights again that even within the same category of exporters, the subset of these which exports from the very beginning is larger on average. As an example application, this implies that some portion of overall firm size distribution changes are driven by compositional changes of when firms begin exporting, a decision which could change due to any number of environmental factors.

Figure 3.2: Firm Size: Exporter Type



Sources: Author's calculations using US Census LFTTD/LBD/CMF as explained in the text.
Notes: This chart shows the kernel density chart of all exporters from the full sample of firms. The blue line shows firms which export from age 1 in the sample and the red line shows firms which export after age 1.

This sample 3.2 of stylized static facts focusing on firm size broken down to according to different dimensions shows the importance of obtaining as much information as possible about exporting firms, as the relationship between exporting and firm size can depend on the type of exporting the firm is engaged in. The next section further highlights the importance of understanding exporting firms in detail, as the dynamics of the relationship between exporting, export intensity, and firm size change throughout the life cycle of an exporter.

3.4 Lifecycle Facts

This section emphasizes the exporter dynamics in the years prior to exit, as well as the probability of observing firm exit. These are traditionally the most understudied aspects of firm life cycle dynamics due to the difficulties of constructing reliable panel data about exporting firms and their domestic establishments. Documenting these dynamics maximally exploits the detailed panel nature of US Census data, especially the sample which tracks firms from birth all the way to firm exit.

Table 3.3: Export Intensity Before Exit

Year Prior to Exit	0	1	2	3	4
observations	500	550	500	550	600
mean	0.097	0.0968	0.086	0.097	0.096
p25	0.0053	0.0056	0.0048	0.0046	0.0048
p75	0.088	0.101	0.081	0.081	0.078
p90	0.257	0.254	0.25	0.279	0.244
std dev	0.193	0.183	0.173	0.196	0.201
median	0.028	0.032	0.022	0.021	0.021

Sources: Author's calculations using US Census LFTTD/LBD/CMF as explained in the text. Balanced Sample
Notes: Export intensity is measured as the ratio of total exports to total value of shipments of firms in each year of the 5 years before exiting from all operations as a firm. Sample numbers are not completely consistent due to fluctuations in the availability of TVS data.

This export intensity before exit table 3.3 shows the trend of export intensity, as measured by the share of total firm sales which are accounted for by exports, in the four years prior to exit. The year of exit is denoted as year 0, and the column title shows how many years before exit the measurement was taken. The source of the total value of shipments variable is the annual survey of manufactures (ASM) and the census of manufactures (CMF) The export intensity trend is relatively stable across the distribution. The sample composition is not completely static due to intermittent surveying of establishments during the years between the CMF, but given that both the mean and the quartiles of the distribution remain stable shows an overall stability in the distribution. The median is slightly increasing, suggesting a faster reduction in domestic sales compared to exports for smaller firms.

The second potential approach to measuring export intensity is to measure the intensity of exports per employee 3.4. The export intensity per employee before exit table shows how even in the balanced sample of this study, firm productivity distributions are quite skewed. The degree to which the mean outpaces even the 75th percentile of firms shows that even the balanced sample of firms that exit gives some generalizable insights about the economy as a whole. While the measure of exported dollars per employee is difficult to interpret in a vacuum, the trends for the distribution as a whole are clear. As firms approach the timepoint at which they exit, export intensity on average increases. Furthermore, the standard deviation of export intensity also increases. This indicates that firm performance as measured by export intensity spreads out, as firms adopt different speeds at which they ramp down employment versus exports.

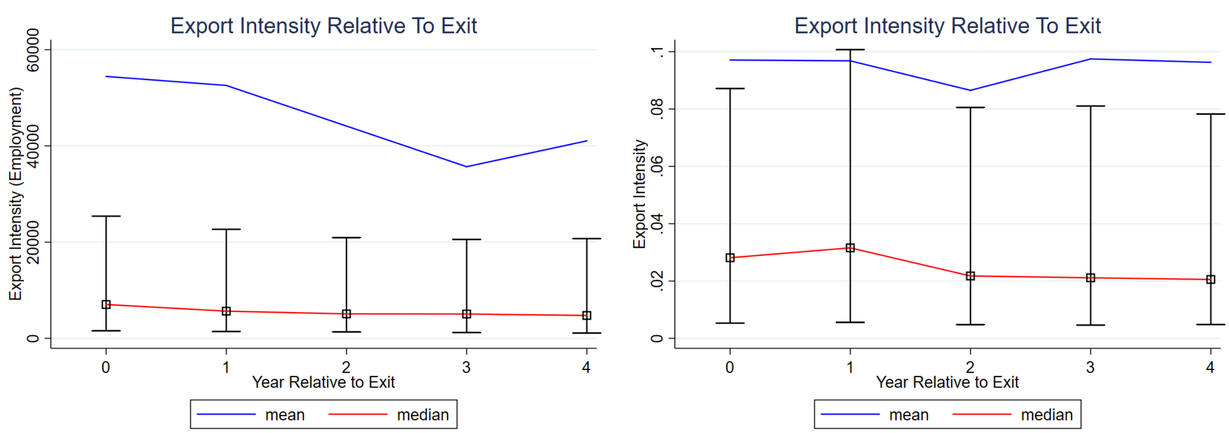
The average export intensity shown in 3.3, as measured by the ratio of exports to the total value of shipments of firms before they exit in the balanced sample is approximately constant over the five years prior to exit. There is a slight increase in the median export intensity in the years prior to exit, indicating that smaller firms are slightly shifting their sales composition to foreign markets. This relatively stable trend further suggests some

Table 3.4: Exports Per Employee Before Exit

Exit year	0	1	2	3	4
observations	1500	1500	1500	1500	1500
mean	54440	52570	44130	35680	41050
p25	1563	1419	1333	1200	1101
p75	25400	22680	20940	20570	20730
p90	72580	73700	68240	70560	66120
std dev	383800	409000	328400	149700	213500
median	7038	5657	5088	5060	4759

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Balanced Sample.
Notes: Export Intensity measured as exports per employee of firms during the five years before exit as a firm. Sample numbers this time remain constant due to the universal coverage of employment data from the LBD.

Figure 3.3: Distribution of export intensity.



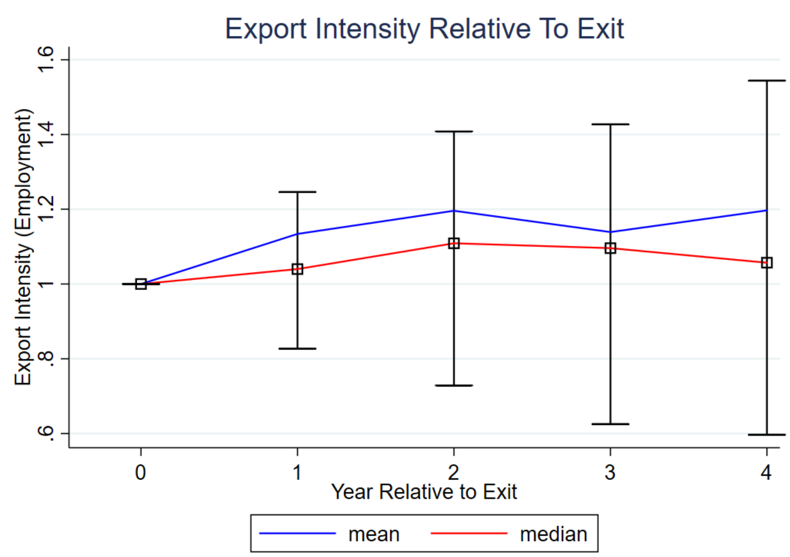
Export Intensity Per Employment

Export Intensity Per Sales

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Balanced Sample.
Notes: The left figure shows export intensity measured in total exports per employee in the five years leading up to firm exit. Year 0 relative to exit is the year we observe the firm ceasing operations, and each year to the right is an additional year before exit. The right figure shows the export intensity measured by total exports divided by total value of shipments (TVS) for the same sample and same definition of five years before exit.

differences in the dynamic behavior of employment versus exports over the life cycle of an exporter, again highlighting the difficulties of imputing exports from employment numbers. The graph shows the mean as a blue line, the median as a red line, and the interquartile range as the black bars. The dramatic difference between the median and the mean highlights that both measures of export intensity are described by a very skewed underlying distribution.

Figure 3.4: Export Intensity Relative To Exit



Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Balanced Sample.
Notes: This chart shows the export intensity, defined as the share of total value of exports out of total value of shipments of firms for which I observe exit in the sample in the four years leading up to exit. This quantity is then divided by its export intensity in its exiting year. The y axis therefore shows relative export intensity in the years leading up to exit. Year 0 relative to exit is the year we observe the firm ceasing operations, and each year to the right is an additional year before exit.

On the other hand, the pattern of export intensity 3.4, measures as total exports per employee, shows that firms ramp up their export intensity in the five years prior to exit. The median firm experiences a slow increase in exports per employee until the final year where

I observe a large increase in exports. The similar position of the mean relative to the 75th percentile indicates that this sample includes very highly productive firms which exit. The increase in the mean follows a similar pattern to the median, ramping up significantly in the year prior to exit and the exit year itself.

Table 3.5 documenting a firm's size in the four years prior to exit relative to its size the year in which it exits using only firms in the balanced sample, which is to say firms which we observe from their birth to exit, further documents significant average life-cycle trends in exporting firms. On average, unsurprisingly, firms shrink in the years leading up to their exit. However, this trend is not universal. The sample contains a number of firms which were growing in the years leading up to their, presumably unexpected, exit. Furthermore, the dispersion in relative sizes is strongly decreasing as firms get closer to their exit year. This compression of the distribution is further suggestive of a steady decline in employment for the average firm, even for higher performers. While we see that the average firm shrinks significantly in total sales in the timeframe before it exits from production, firms around and below the 25th percentile are still growing even in the years leading up to their exit. The average size of a firm that meets the criteria of the sample five years before exit is 2.5 times as large as when it exits. A firm at the 25th percentile five years before exit is only at 80% of its size at exit. Note that these large relative declines could be driven by a relatively small number of absolute employees.

This regression table 3.6 gives us another way to interpret the size findings. This table shows the interaction of the effect of age before exit and size during the life-cycle for the balanced sample of firms. By successively adding the employment control and interactions we

Table 3.5: Size Relative To Exit (Employment)

years before exit	0	1	2	3	4
observations	8500	8500	8500	8500	8500
mean	1	1.671	1.934	2.149	2.485
p25	1	1	0.9286	0.8889	0.7895
p75	1	1.5	1.818	2	2
p90	1	2.444	3	3.5	4
std dev	0	5.843	6.464	7.798	21.83

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Balanced Sample
Notes: This table shows firm size measured by employment in a given year prior to exit relative to firm size at exit. Firm exit happens in year 0, the time variable increases for each year between exit and the measurement. The employment in a given year is then divided by employment in the exit year to give the relative size in each year before exit.

can see if the employment dynamics of a firm prior to exit depend on firm size at a certain age, which seems to not be the case. This analysis with the firm as well as the industry year fixed effect controls for the average size of the firm. However, in the final specification which drops the firm fixed effect, we can see that age at first export is actually a significant predictor of a firm's employment path in the years before it exits. This coefficient implies that firms which were older when they started exporting, which we know from the previous results above tend to be smaller, also shrink slower prior to exiting.

This export intensity by age table 3.7 allows us to disentangle the aggregate statistics reported in the previous one. Again, the measure of export intensity is the exports per employee in the balanced sample. The decreasing negative coefficients on the age variable indicate that, controlling for average employment within a firm, firms are reducing the amount of exports per employee in the years leading up to exit. However, this firm fixed effect

Table 3.6: Balanced Sample Firm Size Relative to Exit

VARIABLES	exit size	exit size	exit size	exit size
age 2	-0.326*	-0.725***	-0.769***	0.396***
	(0.189)	(0.234)	(0.201)	(0.126)
age 3	-0.0134	-0.486***	-0.503***	0.612***
	(0.105)	(0.159)	(0.184)	(0.135)
age 4	-0.00894	-0.350**	-0.432*	0.245
	(0.119)	(0.150)	(0.234)	(0.176)
age 5				0.111
				(0.157)
firm emp		1.614***	1.616***	0.294**
		(0.259)	(0.299)	(0.116)
age 2 * emp			0.0128	-0.296**
			(0.0671)	(0.140)
age 3 * emp			-0.00451	-0.229**
			(0.0550)	(0.116)
age 4 * emp			0.0173	-0.0532
			(0.0599)	(0.105)
age 5 * emp			-0.0236	0.0372
			(0.112)	(0.126)
age at first export				0.0582**
				(0.0250)
Observations	8,500	8,500	8,500	8,500
R-squared	0.657	0.735	0.735	0.438
FE	Firm Ind+Year	Firm Ind+Year	Firm Ind+Year	Ind+Year
clustered	firm	firm	firm	firm

Robust
SE in parentheses
***

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Balanced Sample
Notes: Firm size measured by employment. Standard errors clustered by firm in parentheses. R-sq is reported including the effect of fixed effects. ***p <0.01. This table shows the effect of firm age, employment, and firm age at first export on the size of a firm at the time it exits, measured by employment of the firm. Age at first export is the firm age during the year we first observe exporting. Since this table uses the balanced sample all firms are observed at birth and in their first year of exporting. Firm age in its first year of existence is 1.

precisely controls for average employment over this timespan. The specification without the firm fixed effect show that the employment dynamics of firms leading up to exit are definitely affecting the exports per employee measure, indicating that the two activies do not wind down at similar rates. It's interesting to note that the years before exit effects here depend not just on the size of the firm in the years leading up to exit, as well as the size of the firm at exit, but also the age of first export. This furthers the trend that there seems to be persistent unobserved factors which drive the export entry decision.

Table 3.7: Export Intensity by Age

VARIABLES	exp intensity 1	exp intensity 2	exp intensity 3	exp intensity 4	exp intensity 5	exp intensity 6	exp intensity 7
age 2	-0.147***	-0.448***	-0.536***	0.514***	0.0355***	-0.0332**	0.346***
	(0.0114)	(0.0253)	(0.0371)	(0.0626)	(0.0112)	(0.0144)	(0.0254)
age 3	-0.00411	-0.279***	-0.213***	0.568***	0.337***	0.242***	0.295***
	(0.0113)	(0.0193)	(0.0305)	(0.0709)	(0.0153)	(0.0164)	(0.0484)
age 4	0.0123*	-0.156***	-0.1000***	0.166	0.523***	0.418***	0.151*
	(0.00698)	(0.0113)	(0.0247)	(0.302)	(0.0164)	(0.0124)	(0.0833)
age 5				-0.00175	0.680***	0.575***	0.00673
				(0.304)	(0.0199)	(0.0136)	(0.0874)
log_emp		1.797***	1.727***	0.275***		0.277***	0.228***
		(0.0983)	(0.0906)	(0.0427)		(0.0263)	(0.0233)
age 2 * emp			0.0939***	-0.262***			-0.217***
			(0.0221)	(0.0456)			(0.0237)
age 3 * emp			0.0376	-0.134***			-0.0218
			(0.0243)	(0.0440)			(0.0364)
age 4 * emp			0.0740***	0.155			0.151***
			(0.0271)	(0.133)			(0.0527)
age 5 * emp			0.134***	0.325**			0.312***
			(0.0300)	(0.138)			(0.0569)
age at first export				0.00985***			
				(0.00246)			
Observations	314,000	314,000	314,000	105,000	314,000	314,000	314,000
R-squared	0.599	0.646	0.646	0.092	0.052	0.061	0.065
FE	Firm Ind+Year	Firm Ind+Year	Firm Ind+Year	Ind+Year	Ind+Year	Ind+Year	Ind+Year
clustered	firm	firm	firm	firm	firm	firm	firm
Robust							

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text. Full sample
Notes: Firm size in manufacturing employees. Export intensity measured by exports per employee. Standard errors clustered by firm in parentheses. R-sq is reported including the effect of fixed effects. ***p <0.01. This table shows the effect of firm age, employment on export intensity of a firm, measured by employment of the firm. Age at first export is the firm age during the year we first observe exporting. Firm age in its first year of existence is 1.

Figure 3.7 shows that the effect of size on employment on the probability of exit is qualitatively similar across different categories of exporters. I consider four different sub-samples of the sample of firms: all firms, all exporters, firms which begin exporting in the first five years of their existence, and firms which only engage in non related party exports for their whole existence. For all of these categories, larger firms are less likely to exit. The effect of firm size is largest for the whole sample of firms, about twice as large as the effect for early exporting firms. This again highlights the heterogeneity across firms which would be unobserved by the econometrician who only has access to employment statistics and seeks to predict exit probabilities for a firm.

Table 3.8: Exit Probability

VARIABLES	exit	exit	exit	exit	exit	exit	exit	exit	exit
sub-sample	all	exporter	early exporter	exporter	AL exporter	early exporter	exporter	AL exporter	early exporter
max age				-0.00355***	-0.00353***	-0.00623***			
				(1.62e-05)	(1.97e-05)	(5.44e-05)			
log emp	-0.0687***	-0.0537***	-0.0364***	0.000667***	-0.00139***	-0.00307***	-0.0501***	-0.0552***	-0.0508***
	(0.000388)	(0.000516)	(0.00129)	(0.000107)	(0.000171)	(0.000316)	(0.000482)	(0.000612)	(0.00121)
early exit				0.131***	0.130***	0.141***			
				(0.00194)	(0.00217)	(0.00220)			
Observations	4,224,000	2,000,000	256,000	2,000,000	1,269,000	256,000	2,000,000	1,269,000	256,000
R-squared	0.226	0.170	0.127	0.048	0.051	0.100	0.210	0.223	0.222
FE	firm	firm	firm	ind year	ind year	ind year	firm ind year	firm ind year	firm ind year
clustered	firm	firm	firm	firm	firm	firm	firm	firm	firm
Robust SE									

Sources: Author's calculations using LFTTD/LBD/CMF as explained in the text.
Notes: Probability of observing exit in total sample. The sub-samples in the table above are defined as follows: All refers to the full unbalanced sample. Exporter refers to any firm which exports in its life. A firm is an early exporter if it exports within the first five years of its existence. A firm is an AL exporter if it only engages in arms-length exports, never related party exports. Standard errors clustered by firm in parentheses. ***p <0.01. R-sq is reported including the effect of fixed effects

3.5 Concluding Remarks

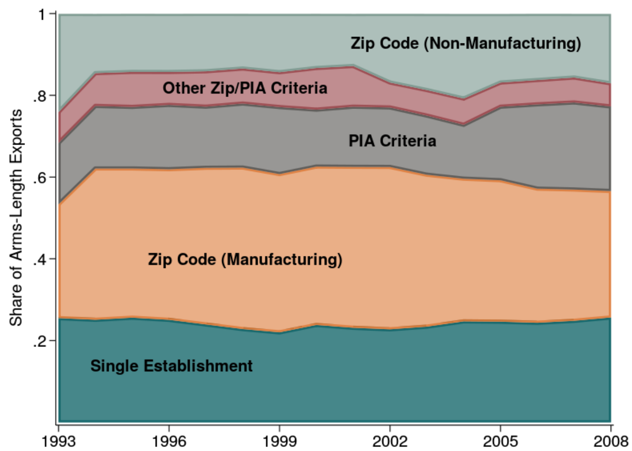
This paper documents static facts and the dynamic behavior of exporting firms in the US Census sample, especially in the period before exit. By observing the whole sample of exporting firms along all stages of their life cycle, I begin to address the question from the introduction: "How does a firm's timing in export decisions affect its life time performance?" I document that exporters appear to differ along in their employment profile their whole life cycle depending on their relationship to exporting at birth. This categorization of different types of exporters also shows up in the static facts. I further show that employment and export dynamics are on average very different across firms in the last few years before their exit. Exporters also differ in how their probability of exit is affected by covariates depending on what kind of exporter they are.

This novel bundle of facts can serve to inform any number of modeling decisions made in economic models of international trade, and shows just one of the many ways in which traditionally unobserved heterogeneity can be alleviated and out of sample predictions about export behavior made more accurate by observing firms across their whole life cycle and better identifying the type of exporters under consideration in a specific scenario. This type of work should have wide applicability to diverse models of trade, and as data availability improves will further improve the tightness of the link between micro-level outcomes and macroeconomic aggregates.

Appendices

.1 Appendix 1

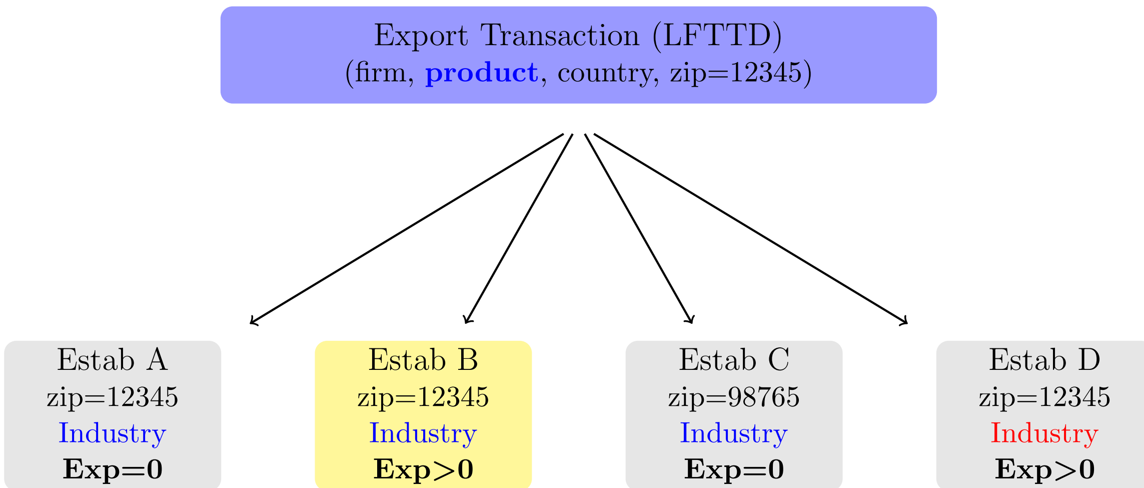
Share of Arms-Length Exports by Method of Allocation



Establishments per Export Transaction

Number of Establishments	Share of Transactions
Arms-Length	
1	81 %
2	8 %
3-5	6 %
6-10	2 %
> 10	3 %
Related-Party	
1	61 %
2	16 %
3-5	12 %
6-10	5 %
> 10	5 %

.



3. Survey-based export indicator
   - Positive exports in ASM or CMF

.

The LBD provides the set of possible establishments associated with firm of transaction

Allocate to Establishments Using 3 Pieces of Information

1. Geographic information
   - Variable in LFTTD identify origin of *movement*, can be different than origin of *production*
2. Production Associated Industry

- Find industries that are principal producers of each product – use Census of Manufacturers Products Trailer File

3. Survey information in ASM/CMF

.2 Aggregation

Aggregating the domestic economy:

To derive the aggregate domestic price level:

$$
\begin{aligned}
P =& \left(\sum_{i=1}^{N} p_d^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P =& N^{\frac{1}{1-\sigma}} \left(\frac{1}{N}\sum_{i=1}^{N} p_d^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P =& N^{\frac{1}{1-\sigma}} \left(\frac{1}{N}\sum_{i=1}^{N} \frac{1}{\rho\phi}^{1-\sigma}\right)^{\frac{1}{1-\sigma}} \\
P =& N^{\frac{1}{1-\sigma}} \frac{1}{\rho} \left(\frac{1}{N}\sum_{i=1}^{N} \phi^{\sigma-1}\right)^{\frac{1}{1-\sigma}} \\
P =& N^{\frac{1}{1-\sigma}} \frac{1}{\rho} \frac{1}{\left(\frac{1}{N}\sum_{i=1}^{N} \phi^{\sigma-1}\right)^{\frac{1}{\sigma-1}}} \\
P =& N^{\frac{1}{1-\sigma}} \frac{1}{\rho\bar{\phi}} = N^{\frac{1}{1-\sigma}} p(\bar{\rho})
\end{aligned}
$$

Aggregate revenue:

$$R_d \equiv \sum_{i=1}^{N} Lp_i c_i$$

Same facts for exporting firms:

$$r_x(\phi) = (P\rho\phi)^{\sigma-1}\tau^{1-\sigma}E\left[R_x\right]$$

$$\pi_d(\phi) = \frac{r_d(\phi)}{\sigma - f}$$

$$\pi_x(\phi) = \frac{r_x(\phi)}{\sigma - f_x}$$

$$f_x = f_{ex} - g(S)$$

.3 HH Problem

Some standard results:

Utility level of household given consumption:

$$C = \left(\sum_{i=1}^{N} c_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$$

HH minimizes expenditure subject to a certain level of utility

$$\min_{\{c_i\}_i^N} \left\{\sum_{i=1}^{N} p_i c_i \quad s.t. \quad \left(\sum_{i=1}^{N} c_i^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}} \geq \bar{C}\right\}$$

FOC:

$$c_i = (\frac{p_i}{P})^{-\sigma} C$$

where P is

$$P = \left(\sum_{i=1}^{N} p_i^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$$

and C is as above.

Also

$$r_i = \left(\frac{p_i}{P} \right)^{1-\sigma} R$$

where $R = \sum_{i=1}^{N} LP_iC_i$

.4 Appendix 2

.4.1 Sources of US Data on Firm and Establishment Level Exports

The original source for firm and establishment level export data from the US Census Bureau comes from the quinquennial Census of Manufacturers (as well as the annual supplement, the Annual Survey of Manufacturers). This survey asks establishments to report the dollar value of their shipments that are destined for foreign countries. The advantage of this survey-based question is the tight mapping between export values and a particular manufacturing plant ostensibly involved with the actual production of that exported product. There are numerous disadvantages of this source, however. There is no product or destination-level detail, it is only an annual measure of total exports, there are concerns of reliability due to the survey basis of the reporting, there is no information on the import side, and the longitudinal nature of the data is limited.

Detailed information on trade transactions are compiled by US Customs for purposes of enforcing trade laws, documenting trade flows, and monitoring border security. The administrative documentation attached to these transactions offers rich detail, including the value, quantity, detailed product, country of export/import, port location, type of transaction, and more. Beginning with Bernard et al. (2009), the US firm associated to these trade transactions was linked to other Census datasets, thus improving the dimensions of information on firm-level trade for study by economists. The principal method of linkage was the employer identification number (EIN), as it was recorded by Customs Bureau documents accompanying individual shipments and also existed as part of the establishment/firm register in the Census Bureau.[6] The resulting Linked/Longitudinal Foreign Trade Transactions (LFTTD) database has been a very useful resource for trade economists studying import/export patterns by US firms. One disadvantage, however, is that the EIN-based matching does not allow exports or imports to be assigned to individual establishments (plants) of the firm. The drawbacks of this limitation are discussed further below.

The final major resources for firm-level trade are the surveys of multinational firms collected by the Bureau of Economic Analysis. The level of aggregation is a mix between the firm and establishment (affiliate), depending on the direction of trade, size of the firm, or whether a particular year falls under the BEA's benchmark survey period. Like the survey-based information from the Census Bureau, the BEA data do not have extensive information on products or high-frequency detail of shipments. And while the focus on multinational firms provides for useful splits between arms-length and related-party trade, the BEA data do not have any information on non-multinational firms.

.4.2 Details on Establishment-Level Allocation Methodology

.4.3 Industry-Based Matching

Every five years as part of the Census of Manufacturers, the Census Bureau surveys establishments on their total shipments broken down into a set of NAICS-based (6 digit) product categories.[7] Each establishment is given a form—specific to its industry—with a list of pre-specified products. There is also additional space to record other product shipments

[6] On the export side, the EIN is listed on the "Shippers Export Declaration. The exception are shipments to Canada, which do not contain EINs but rather a field listing the firm name. On the import side, the Customs Forms (7501 and 7503) record the EIN representing the "ultimate consignee" of the imported goods.

[7] The 1992 version of the CM used SIC-based product codes.

not included in the form. The resulting product trailer file to the CM allows a researcher to construct the set of industries that are primary producers of a given product.

There are several data issues that must be addressed before using the CM-Products file to infer information about the relative value of product-level shipments by a particular firm. First, the trailer file contains product-codes that are used to "balance" the aggregated product-level value of shipments with the total value of shipments reported on the base CM survey form. We drop these product codes from the dataset. Second, there are often codes that do not correspond to any official 7-digit product code identified by Census. (These are typically products that are self-identified by the firm but do not match any of the pre-specified products identified for that industry by Census.) Rather than ignoring the value of shipments corresponding to these codes, we attempt to match at a more aggregated level. Through an iterated process we try to find a product code match at the 6, 5, and 4 digit product code level, and use the existing set of 7-digit matches as weights to allocate the product value among the 7-digit product codes encompassed by the more aggregated level.

Finally, the link between the Harmonized Commodity Description and Coding System (or Harmonized System, HS) codes and Standard Industrial Classification System (SIC) and North American Industrial Classification System (NAICS) product codes is referred to as a SIC base or NAICS base, depending on which CM year is being used. These basecodes are up to 8 alphanumeric characters long, with shorter basecodes representing more highly aggregated products. Given linkage between either SIC or NAICS, the first four to six digits of the basecodes are called the baseroot. Each HS code has a single baseroot, while a baseroot might be associated with multiple HS codes. We use the NAICS (or SIC) to HS concordance from Pierce and Schott (2012), to map the information from the CM-Products file to the

LFTTD trade data.

We now describe how we construct the set of "Production-Associated Industries (PAIs)" associated with a given product. Formally, let x_{pij} denote the value of shipments of product p by establishment i in industry j during a census year. Then the total output of product p in industry j can be written as:

$$X_{pj} = \sum_{i=1}^{I_j} x_{pij},$$

where I_{jp} is the number of establishments producing p in industry j. Total output of product p is then:

$$X_p = \sum_{j=1}^{I_{jp}} X_{pj}.$$

The share of product output accounted for by a given industry j is therefore:

$$S_{pj} = \frac{X_{pj}}{X_p}.$$

Because of reporting errors and aggregation of products, we designate an industry as a PAI of product p provided that its share S_{pj} passes a certain threshold – which we set at 5 percent [8]. We define the set of industries for product p for which $S_{pj} > 0.05$ as J_p.

We match individual years of the LFTTD data to the closest available Census year. To summarize, this procedure allows one to attach a set of industries to each exported product from the LFTTD; the industry detail of establishments in the LBD can then be used to match products to establishments.

[8]We have varied this threshold without affecting our results

.4.4 Location-Based Matching

Along with industry-based information on production, we also use geographic information to narrow the set of potential establishments involved with a particular trade transaction. Part of the shippers export declaration form (now electronically administered via the Automated Exporter System) asks for address information on the USPPI where "goods begin their journey to the Port of Export." Both the zipcode and state information from this entry are included in the LFTTD microdata.

Although uncommon, some trade transactions in the LFTTD record a missing or incomplete zipcode. For these observations, we fill in missing zip-codes iteratively by replacing missing values with the largest zip-code value within the firm, country, month, HS code and baseroot observation. By attaching a US location to an export transaction, this method may also assist in identifying the relevant establishment of export, though subject to the limitations outlined in the main text. We will discuss below the issue of whether these establishments are the location of production or the location of export processing.

.4.5 Survey-Based Information

A final set of information used to allocate firm-level exports to individual establishments is the export variable included in the CMF and ASM. Although the CMF should be comprehensive across all manufacturing establishments, one must be more careful in ASM years given that not all of a firm's manufacturing plants may be included in the survey sample. For this reason we use the export indicator from the CMF/ASM as a way of distinguishing between a firm's plants when the industry/location information specified above yields multiple plants associated with a given transaction.

.4.6 Allocation Procedure

The paragraphs below outline how we combine these sources of information to allocate all firm-level LFTTD exports to the most likely establishment associated with that export. This allocation will not always identify the establishment of export manufacture for several reasons. First, the PPI identified in the export declaration may be a non-manufacturing firm entirely that is solely involved in the export of the good; in such cases it is impossible to identify the establishment of production. Second, if the PPI is the firm of manufacture but processes shipments for export in a separate establishment, then our location information will point to a non-production establishment. Hence, while the allocation procedure described below attempts to prioritize establishments of production over non-manufacturing establishments, the data will often only identify the establishment involved with the export process. It is worth emphasizing that both production and non-production establishments involved with exporting activity will be impacted by trade and export markets.

To retain as much detail as possible, we take the raw LFTTD export data and aggregate only up to the firm, product, country, month, zipcode, port, and export-method (rail, air, etc) level. Next we take the cartesian product of these firm-level transactions to the full set of firm-establishments from the LBD in each associated year – essentially making copies of each export observation to attach to all of the firm's establishments. Because the LBD only registers a firm if it existed on March 12th of a particular year, some firms could be trading in the LFTTD but not exist in the LBD. To remedy this issue, we match the trade data not found in the LBD for that period with samples from the year prior and the year following. Using this large dataset, we retain the most likely establishment for each trade observation according to an iterative set of rules, decreasing in the degree of confidence in the establishment match.

- **Case 1: Single unit firm.** The allocation is a trivial exercise for those firms having only one establishment. We first remove these transactions, but flag whether these establishments are manufacturing or non-manufacturing, and whether the establishment records positive export shipments in the ASM/CM.

- **Case 2: Unique zip code match: manufacturing.** If a single zip code matches to a unique establishment, then the relevant trade is assigned to that establishment. We separately flag whether the establishment records positive export shipments in the ASM/CM.

- **Case 3: Non-unique zip code match to PAI establishment** If there is only one establishment matching the zip code that also matches based on our PAI criteria, then we allocate all trade to the zipcode match that also aligns with the appropriate industry. Given that an HS code is assigned to multiple PAIs, it is possible for there to be several establishments matching this case. Absent any distinguishing information on export activity from the ASM/CM, we use employment weights to allocate the exports across establishments.

- At this stage for any unallocated export transactions, we loop back through cases 3 through 6 but looking for PAI establishments at the NAICS-5, and then NAICS-4 industry basis.

- **Case 4: No zip code match but unique PAI establishment.** In this case, simply allocate all trade to the unique PAI establishment.

- **Case 5: No zip code match but non-unique PAI establishments.** For multiple establishments matching a PAI, we must use employment weights to allocate the trade across matching establishments (absent distinguishing information on export status from the ASM/CM).

- **Case 6: Non-unique zip code match, and no PAI establishments.** We split these cases into those matching manufacturing establishments with those matching non-manufacturing establishments. If there are multiple establishments with the same zip code (which is rare), we continue to use employment shares as weights.

- **Case 7: No zip code match and no PAI establishments.** For this final case, we first allocate the export transaction to all manufacturing establishments of the associated firm (using either ASM/CM export share weights or employment weights). We also prioritize establishments in the wholesale (NAICS 42) and transportation/warehousing (NAICS 48) industries, based on the distribution of exports from prior matches. if there are no establishments in any of these industries, then the final step allocates the export transaction to all other establishments based on employment shares.[9]

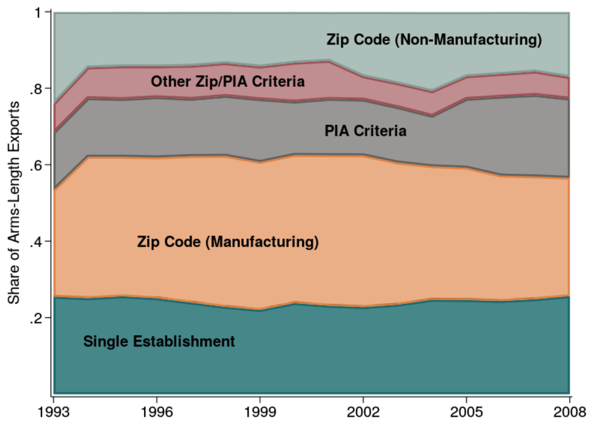
.4.7 Characteristics of the Allocation

Figure 5 documents the share of overall US exports that is allocated according to the hierarchy of methods as described above.

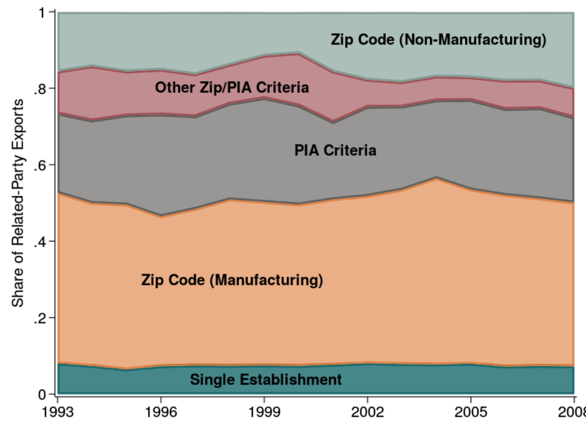
[9]In reality, for these cases we take the top 20 establishments by employment size, as there are some firms with thousands of establishments.

Figure 5: Share of Exports by Type of Establishment Allocation

(a) Non-Related Party Exports



(b) Related Party Exports



Source: Author's calculations as explained in text.

.4.8 Additional Results

.4.9 Public vs Allocated

Table 9: Economic Activity HHI Index

employment	pay	Total Trade	TVS	year
0.004	0.007	0.020	0.057	2007
0.004	0.006	0.020	0.004	2009

.4.10 Leverage Regression Tables

As an additional robustness check, we control explicitly for domestic (US) demand in all of our regression specifications. We build on Stumpner (2019) in the constructing a leverage based domestic demand shock. This approach has the advantage of taking into account inter-county domestic trade links in the US when measuring domestic demand.

$$TDS_t^k = \sum_{n-1}^{N} \frac{X_{nt}^k}{Y_t^k} Lev_n \tag{1}$$

The specification of the domestic demand shock we disclose uses a 3 digit NAICS level of industry aggregation, but the sign and significance of the results are not sensitive to the choice of industry aggregation level nor to the specific measure of housing price drops or county leverage increases.

.4.11 Concentration Facts

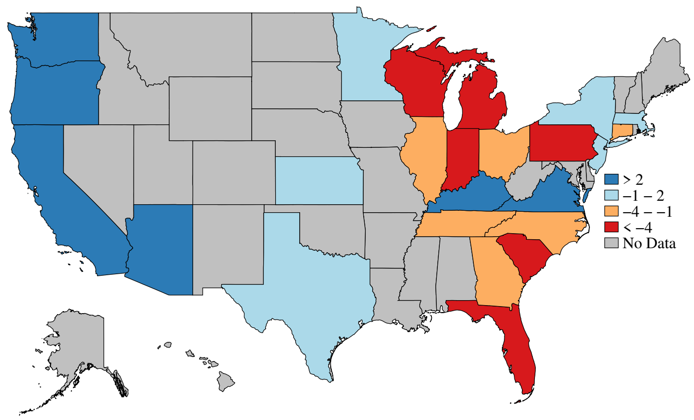
Table 10: Public Use Data Vs Allocated Data

State	Share in public data (in percent)	Share in allocation (in percent)	Percentage point difference
Texas	16.3	10.3	6.0
Florida	4.2	3.3	0.9
Louisiana	1.8	1.3	0.5
Kentucky	1.9	1.7	0.3
Washington	4.5	4.2	0.3
Georgia	2.2	1.9	0.3
Vermont	0.4	0.1	0.3
Maryland	0.8	0.6	0.2
Iowa	0.9	0.7	0.2
Utah	0.7	0.5	0.2
Nevada	0.5	0.4	0.2
Idaho	0.5	0.3	0.1
Arizona	1.8	1.7	0.1
NewMexico	0.3	0.2	0.1
Virginia	1.5	1.4	0.1
SouthDakota	0.1	0.1	0.1
WestVirginia	0.3	0.3	0.1
NorthDakota	0.2	0.1	0.1
Wyoming	0.1	0.0	0.0
Tennessee	2.1	2.1	0.0
Alabama	1.3	1.3	0.0
Maine	0.2	0.2	0.0
Alaska	0.1	0.0	0.0
Kansas	0.9	0.9	0.0
Arkansas	0.5	0.5	0.0
Mississippi	0.5	0.5	0.0
Delaware	0.4	0.4	0.0
Oklahoma	0.4	0.5	-0.1
Nebraska	0.4	0.4	-0.1
NewHampshire	0.3	0.3	-0.1
RhodeIsland	0.1	0.2	-0.1
Massachusetts	2.4	2.6	-0.2
Oregon	1.4	1.6	-0.2
SouthCarolina	1.6	1.9	-0.2
Montana	0.1	0.3	-0.2
Ohio	4.2	4.4	-0.3
Wisconsin	1.8	2.1	-0.3
NewJersey	2.9	3.1	-0.3
Colorado	0.7	1.0	-0.3
Pennsylvania	2.8	3.1	-0.3
NorthCarolina	2.2	2.6	-0.4
NewYork	5.9	6.3	-0.4
Indiana	2.6	3.0	-0.4
Connecticut	1.3	1.9	-0.5
Illinois	4.7	5.2	-0.5
Minnesota	1.7	2.2	-0.6
Missouri	1.2	1.9	-0.7
California	12.3	13.2	-0.9
Michigan	4.2	7.0	-2.8

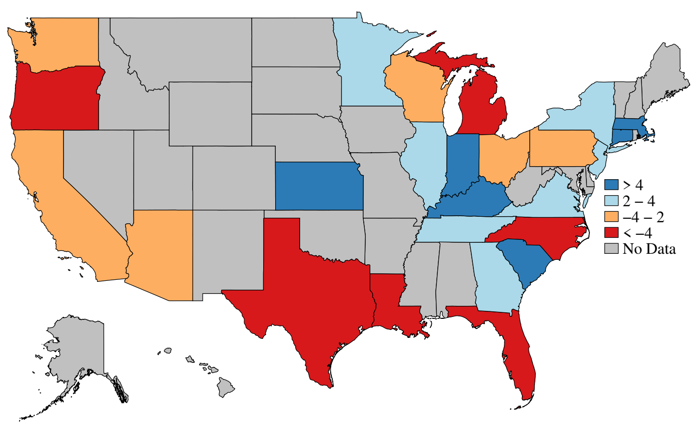
Sources: Author's calculations using LFTTD, CMF as explained in the text. Census Bureau State-Level Trade
Notes: This chart shows the share of trade each state accounts for in its respective data-source in 2007, as well as the percentage difference between the two. A positive percent difference indicates that the public use data attributes more exports to a specific state than the establishment allocation
Hawaii is excluded for disclosure purposes

Figure 6: State-Level Exports to Select Regions, Percent of Total

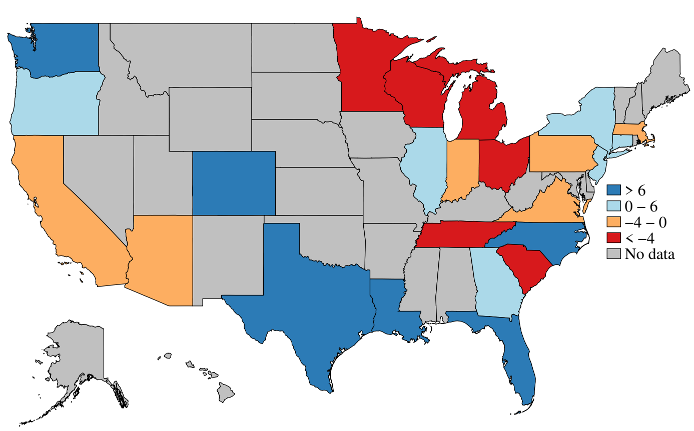
(a) State-Level Exports to Asia (excl. China)



(b) State-Level Exports to Europe



(c) State-Level Exports to Rest of World



Sources: Author's calculations using LFTTD, CMF as explained in the text.
Notes: Europe is defined as... Rest of World excludes Europe, Canada, Mexico, and Asia (including China).

Table 11: House Leverage Effects

Variable	Full Sample			Manufacturing			Non-manufacturing		
	Emp	Pay	Wage	Emp	Pay	Wage	Emp	Pay	Wage
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	3.18 **	0.275 ***	0.755 ***	1.45 **	0.177 ***	1.34 ***	2.11	0.0654 *	0.268 *
	(1.30)	(0.0488)	(0.130)	(0.587)	(0.0286)	(0.280)	(1.15)	(0.0377)	(0.146)
lagged dependent variable	-0.07 ***	-0.08 ***	-0.3 ***	-0.005	-0.14 ***	-0.29 ***	-0.11 ***	-0.71 ***	-0.125 ***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.04)	(0.13)	(0.03)
$s_{c,2007}^{\text{emp,exp}}$	0.00	0.02	0.00	0.00	0.03	0.00	0.00	-0.04	0.00
	(0.00)	(0.04)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.03)	(0.00)
house price leverage	-0.021 ***	-0.856 ***	-0.014 *	-0.002	-0.054	-0.018	-0.02 **	-0.65 *	-0.019 **
	(0.01)	(0.22)	(0.01)	(0.00)	(0.08)	(0.02)	(0.01)	(0.35)	(0.01)
State fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	2600	2600	2600	2600	2600	2600	2600	2600	2600

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports.
In this table each column contains the same regression specification, using the full set of controls and state fixed effects. The first three columns display estimation results for total county level employment, pay, and wage responses. The second set of three columns displays results for the responses of manufacturing dependent variables, and the last set of three columns shows results for the responses of non-manufacturing dependent variables.
In all specifications the employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. *Lagged dependent variable* refers to the dependent variable from 2005 to 2007. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table 12: House Leverage Effects with Neighbor

Variable	Full Sample			Manufacturing			Non-manufacturing		
	Emp	Pay	Wage	Emp	Pay	Wage	Emp	Pay	Wage
$\frac{\text{exp}_{c,2009}-\text{exp}_{c,2007}}{\text{emp}_{c,2007}}$	3.19 **	0.275 ***	0.758 ***	1.45 **	0.18 ***	1.5 ***	2.12 *	0.07 *	0.271 *
	(1.30)	(0.0488)	(0.131)	(0.587)	(0.0286)	(0.416)	(1.15)	(0.0377)	(0.147)
neighbor county trade growth	0.002	0.15 **	0.004 **	0.00	0.04	0.01 **	0.00	0.1	0.41 *
	(0.00209)	(0.0590)	(0.00197)	(0.000769)	(0.0309)	(0.00354)	(0.00240)	(0.0665)	(0.00226)
lagged dependent variable	-0.076 **	-0.086 ***	-0.301 ***	-0.006	-0.142 ***	-0.284 ***	-0.107 **	-0.707 ***	-0.125 ***
	(0.028)	(0.025)	(0.028)	(0.034)	(0.033)	(0.03)	(0.042)	(0.13)	(0.03)
$s_{c,2007}^{\text{emp,exp}}$	0.00	0.03	0.00	0.00	0.03	0.00	0.00	-0.03	0.00
	(0.00)	(0.04)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.03)	(0.00)
house price leverage	-0.021 **	-0.81 ***	-0.013 *	-0.002	-0.042	-0.013	-0.019 **	-0.62 *	-0.017 **
	(0.00801)	(0.224)	(0.00742)	(0.00220)	(0.0801)	(0.0125)	(0.00957)	(0.343)	(0.00878)
State fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	2600	2600	2600	2600	2600	2600	2600	2600	2600

Notes: The table displays two-stage least squares coefficient estimates based on equation (2.2), where we use the WID instrument (2.4) to address the endogeneity of exports.
In this table each column contains the same regression specification, using the full set of controls, the neighbor county spillover measure and state fixed effects. The first three columns display estimation results for total county level employment, pay, and wage responses. The second set of three columns displays results for the responses of manufacturing dependent variables, and the last set of three columns shows results for the responses of non-manufacturing dependent variables.
In all specifications the employment share in exporting establishments is constructed according to equation (2.3) and measured in percent. *Lagged dependent variable* refers to the dependent variable from 2005 to 2007. The number of observations—rounded to the closest 100 to comply with Census disclosure requirements—is 2600 in all specifications. Bootstrapped standard errors are reported in parentheses. ***, **, and * indicate significance at the 1, 5, and 10 percent level.

Table 13: Lorenz Curves

half-decile	employment share	pay share	county tvs share	total trade share
6	0.01303	0.00768	0.00248	0.00048
7	0.01786	0.01074	0.00458	0.00093
8	0.02379	0.01466	0.00761	0.00168
9	0.03103	0.01947	0.01188	0.00290
10	0.03974	0.02532	0.01770	0.00476
11	0.05016	0.03248	0.02556	0.00767
12	0.06296	0.04133	0.03636	0.01196
13	0.07883	0.05249	0.05027	0.01828
14	0.09820	0.06626	0.06817	0.02760
15	0.12250	0.08438	0.09137	0.04139
16	0.15580	0.10990	0.12310	0.06237
17	0.20430	0.14710	0.16690	0.09504
18	0.27670	0.20530	0.23210	0.15260
19	0.41980	0.33100	0.34980	0.28170
20	1.00000	1.00000	1.00000	1.00000

Sources: US Census; authors' calculations.

Notes: Table displays the cumulative sums of the column variable at a national level by counties sorted into half-deciles (20 bins) for each variable.
Bottom 5 half-deciles suppressed due to disclosure protocol

Table 14: Economic Activity HHI Index

employment	pay	total trade	tvs	year
0.004	0.007	0.020	0.057	2007
0.004	0.006	0.020	0.004	2009

Sources: US Census; authors' calculations.

Notes: Table displays the Herfindahl–Hirschman Index for each variable in 2007 and 2009 Bottom 5 half-deciles suppressed due to disclosure protocols

.5 Appendix 3

Figure 7: Cumulative Exports

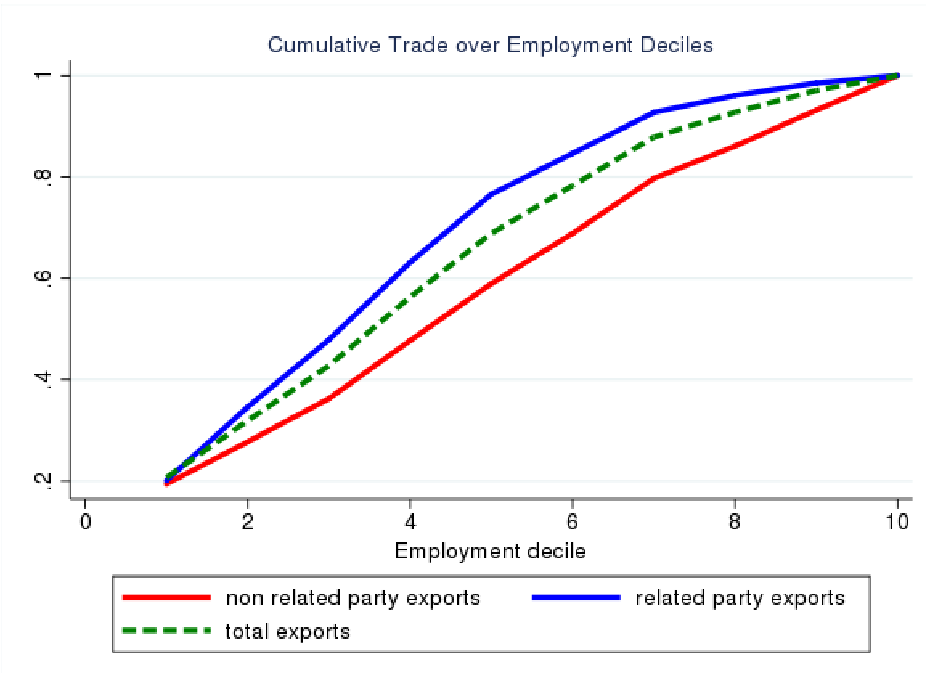
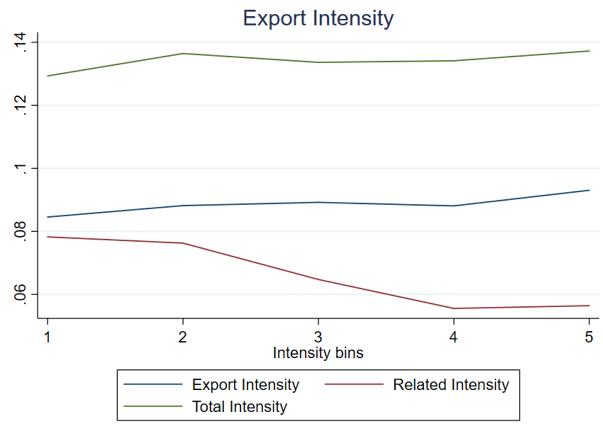


Table 15: Export Participation by Year and Bin

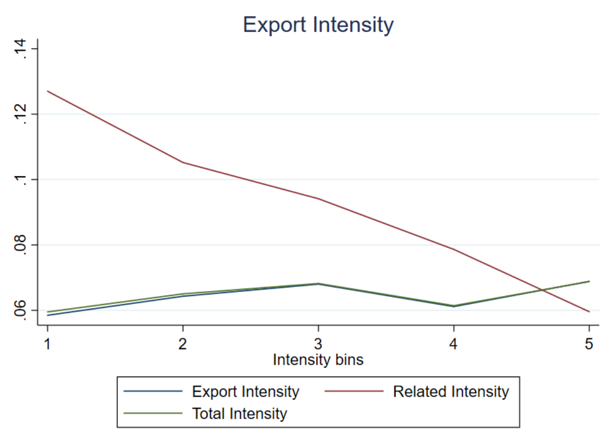
Year	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	
Export Participation	0.3698	0.3828	0.3951	0.4073	0.422	0.4349	0.4456	0.4551	0.4644	0.4715	0.4804	0.4871	0.4929	0.5001	0.5012	0.5058	0.5118	0.5166	0.5227	0.5248	0.5287	0.5315	Bin 1
	0.9504	0.9523	0.9557	0.9599	0.9614	0.9624	0.9662	0.9642	0.969	0.9691	0.9694	0.9682	0.9693	0.9695	0.9667	0.9664	0.9665	0.9726	0.9737	0.9714	0.9704	0.9688	Bin 2
	0.9898	0.9917	0.9926	0.9936	0.9944	0.9963	0.9962	0.9961	0.997	0.996	0.9949	0.9959	0.9958	0.9945	0.9955	0.9954	0.9953	0.9988	0.9975	0.9988	0.9987	0.9948	Bin 3

Figure 8: Export Intensity (Full Sample)



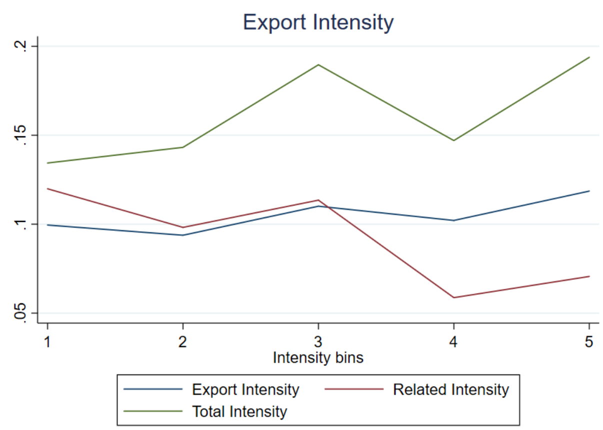
Export Intensity by Firm Age Full Sample

Figure 9: Export Intensity (Pure Exporters)



Export Intensity by Firm Age Pure Exporters

Figure 10: Export Intensity (Balanced Sample)



Export Intensity by Firm Age Balanced Sample

Bibliography

Ahn, J., M. Amiti, and D. E. Weinstein (2011, May). Trade Finance and the Great Trade Collapse. *American Economic Review 101*(3), 298–302.

Aitken, B., G. Hanson, and A. Harrison (1997). Spillovers, foreign investment, and export behavior. *Journal of International Economics 43*(1-2), 103–132.

Albornoz, F., H. F. C. Pardo, G. Corcos, and E. Ornelas (2012). Sequential exporting. *Journal of International Economics 88*(1), 17 – 31.

Alessandria, G. and H. Choi (2014). Establishment heterogeneity, exporter dynamics, and the effects of trade liberalization. *Journal of International Economics 94*(2), 207–223.

Alessandria, G., J. P. Kaboski, and V. Midrigan (2010a, December). Inventories, lumpy trade, and large devaluations. *American Economic Review 100*(5), 2304–39.

Alessandria, G., J. P. Kaboski, and V. Midrigan (2010b, December). The Great Trade Collapse of 2008–09: An Inventory Adjustment?. *IMF Economic Review 58*(2), 254–294.

Allen, T. and C. Arkolakis (2014). Trade and the topography of the spatial economy. *Quarterly Journal of Economics 129*(3), 1085–1140–48.

Atalay, E., A. Hortaçsu, and C. Syverson (2014, April). Vertical integration and input flows. *American Economic Review 104*(4), 1120–48.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American Economic Review 103*(6), 2121–68.

Axtell, R. (2001). Zipf distribution of u.s. firm sizes. *Science 293*, 1818–1820.

Bayard, K., D. Byrne, and D. Smith (2015). The scope of factoryless manufacturing. In S. N. Houseman and M. Mandel. (Eds.), *Measuring Globalization: Better Trade Statistics for Better Policy – Volume 2. Factoryless Manufacturing, Global Supply Chains, and Trade in Intangibles and Data.*, pp. 81–120. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Bems, R., R. C. Johnson, and K.-M. Yi (2013, May). The Great Trade Collapse. *Annual Review of Economics 5*(1), 375–400.

Benguria, F. and F. Saffie (2019). Dissecting the Impact of the 2018-2019 Trade War on U.S. Exports. *SSRN Electronic Journal*.

Benguria, F. and F. Saffie (2020). The Impact of the 2018-2019 Trade War on U.S. Local Labor Markets. *SSRN Electronic Journal*.

Bernard, A., B. Jensen, and P. Schott (2009). Importers, exporters, and multinationals: A portrait of firms in the u.s. that trade goods. In T. Dunne, J. B. Jensen, and M. J. Roberts (Eds.), *Producer Dynamics: New Evidence from Micro Data*, pp. 513–552. Chicago: University of Chicago Press.

Bernard, A., J. B. Jensen, S. J. Redding, and P. K. Schott (2007). Global firms. *Journal of Economic Literature 56*(2), 565–619.

Bernard, A. B., E. A. Boler, R. Massari, J.-D. Reyes, and D. Taglioni (2017, October). Exporter Dynamics and Partial-Year Effects. *American Economic Review 107*(10), 3211–28.

Bernard, A. B. and T. C. Fort (2015, May). Factoryless Goods Producing Firms. *American Economic Review Papers and Proceedings 105*(5), 518–523.

Bernard, A. B. and J. B. Jensen (1999). Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics 47*, 1–25.

Bernard, A. B. and J. B. Jensen (2004). Why Some Firms Export. *The Review of Economics and Statistics 86*(2), 561–569.

Besedeš, T. and T. J. Prusa (2011, November). The role of extensive and intensive margins and export growth. *Journal of Development Economics 96*(2), 371–379.

Boehm, C. and N. Pandalai-Nayar (2020, March). Convex Supply Curves. NBER Working Papers 26829, National Bureau of Economic Research, Inc.

Boehm, C., N. Pandalai-Nayar, A. Flaaen, and J. Schlupp. The Geography of Exports.

Borusyak, K., P. Hull, and X. Jaravel (Forthcoming). Quasi-experimental shift-share research designs. *Review of Economic Studies*.

Bown, C. P. and M. A. Crowley (2013). Import protection, business cycles, and exchange rates: Evidence from the Great Recession. *Journal of International Economics 90*(1), 50–64.

Bussière, M., G. Callegari, F. Ghironi, G. Sestieri, and N. Yamano (2013, July). Estimating Trade Elasticities: Demand Composition and the Trade Collapse of 2008-2009. *American Economic Journal: Macroeconomics 5*(3), 118–151.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock. *Econometrica 87*(3), 741–835.

Caliendo, L., F. Parro, E. Rossi-Hansberg, and P.-D. Sarte (2018, October). The Impact of Regional and Sectoral Productivity Changes on the U.S. Economy. *The Review of Economic Studies 85*(4), 2042–2096.

Cassey, A. (2009). State Export Data: Origin of Movement vs. Origin of Production. *Journal of Economic and Social Measurement 34*, 241–268.

Cassey, A. and K. Schmeiser (2013). The Agglomeration by Destination of U.S. State Exports. *Economics Bulletin 33*, 1504–1510.

Castillo-Giménez, J., G. Serrano, and F. Requena-Silvente (2011, January). Exports dynamics and information spillovers: evidence from Spanish firms. Working Papers 1103, Department of Applied Economics II, Universidad de Valencia.

Census, U. (2008, February). FT-900 SUPPLEMENT December 2007. Technical Report Exhibit 1, US Census.

Choquette, E. and P. Meinen (2015). Export Spillovers: Opening the Black Box. *The World Economy 38*(12), 1912–1946.

Coibion, O. and Y. Gorodnichenko (2015). Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts. *The American economic review 105*(8), 2644–2678.

Cragg, J. G. and S. G. Donald (1993, April). Testing Identifiability and Specification in Instrumental Variable Models. *Econometric Theory 9*(2), 222–240.

Crozet, M. and P. Koenig (2010). Structural gravity equations with intensive and extensive margins. *Canadian Journal of Economics 43*(1), 41–62.

Deng, Z., H. Guo, W. Zhang, and C. Wang (2014). Innovation and survival of exporters: A contingency perspective. *International Business Review 23*(2), 396–406.

Deng, Z., R.-J. B. Jean, and R. R. Sinkovics (2017, April). Polarizing Effects of Early Exporting on Exit. *Management International Review 57*(2), 243–275.

Dix-Carneiro, R. and B. K. Kovak (2017, October). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–46.

Eaton, J., S. Kortum, B. Neiman, and J. Romalis (2016, November). Trade and the global recession. *American Economic Review 106*(11), 3401–38.

Feenstra, R. C., Z. Li, and M. Yu (2014, October). Exports and Credit Constraints under Incomplete Information: Theory and Evidence from China. *The Review of Economics and Statistics 96*(4), 729–744.

Fernandes, A. and H. Tang (2014). Learning to Export from Neighbors. pp. 45.

Gabaix, X. (1999, 1999). Zipf's law for cities: An explanation. *Quarterly Journal of Economics 114*(3), 739–767.

Gabaix, X. (2011). The granular origins of aggregate fluctuations. *Econometrica 79*, 733–772.

Greenaway, S. D., N. Sousa, and K. Wakelin (2004). Do domestic firms learn to export from multinationals? *European Journal of Political Economy 20*(4), 1027–1043.

Gumpert, A., H. Li, A. Moxnes, N. Ramondo, and F. Tintelnot (2020). The life-cycle dynamics of exporters and multinational firms. *Journal of International Economics 126*, 103343.

Hakobyan, S. and J. McLaren (2016, 10). Looking for Local Labor Market Effects of NAFTA. *The Review of Economics and Statistics 98*(4), 728–741.

House, C. L., C. Proebsting, and L. L. Tesar (2019, July). Regional effects of exchange rate fluctuations. Working Paper 26071, National Bureau of Economic Research.

Hummels, D., R. J. rgensen, J. Munch, and C. Xiang (2014). The wage effects of offshoring: Evidence from danish matched worker-firm data. *American Economic Review 104*(6), 1597–1629.

Ilmakunnas, P. and S. Nurmi (2010). Dynamics of Export Market Entry and Exit*. *The Scandinavian Journal of Economics 112*(1), 101–126. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9442.2009.01596.x.

Impullitti, G., A. A. Irarrazabal, and L. D. Opromolla (2013). A theory of entry into and exit from export markets. *Journal of International Economics 90*(1), 75–90.

Jarmin, R. and J. Miranda (2002). The longitudinal business database. mimeo, available at http://www.vrdc.cornell.edu/info7470/2007/Readings/jarmin-miranda-2002.pdf.

Jovanovic, B. (1982). Selection and the Evolution of Industry. *Econometrica 50*(3), 649–670.

Koenig, P., F. Mayneris, and S. Poncet (2010, May). Local export spillovers in France. *European Economic Review 54*(4), 622–641.

Kovak, B. K. (2010, July). Regional Labor Market Effects of Trade Policy: Evidence from Brazilian Liberalization. Working Papers 605, Research Seminar in International Economics, University of Michigan.

Levchenko, A. A., L. T. Lewis, and L. L. Tesar (2010, December). The Collapse of International Trade during the 2008–09 Crisis: In Search of the Smoking Gun. *IMF Economic Review 58*(2), 214–253.

Lincoln, W. F. and A. H. McCallum (2018, February). The rise of exporting by U.S. firms. *European Economic Review 102*, 280–297.

Melitz, M. J. (2003, November). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica 71*(6), 1695–1725.

Muûls, M. (2008, September). Exporters and credit constraints. A firm-level approach. Working Paper Research 139, National Bank of Belgium.

Pierce, J. and P. Schott (2012). A concordance between u.s. harmonized system codes and sic/naics product classes and industries. *Journal of Economic and Social Measurement 37*(1-2), 61–96.

Ruhl, K. J. and J. L. Willis (2017). New exporter dynamics. *International Economic Review 58*(3), 703–726.

Sanderson, E. and F. Windmeijer (2016). A weak instrument F-test in linear IV models with multiple endogenous variables. *Journal of Econometrics 190*(2), 212–221.

Santamaría, M., J. Ventura, and U. Yeşilbayraktar (2020, December). Borders within Europe. Technical Report w28301, National Bureau of Economic Research, Cambridge, MA.

Sjöholm, F. (1999). Do Foreign Contacts Enable Firms to Become Exporters?

Stock, J. and M. Yogo (2005). *Testing for Weak Instruments in Linear IV Regression*, pp. 80–108. New York: Cambridge University Press.

Stumpner, S. (2019, July). Trade and the geographic spread of the great recession. *Journal of International Economics 119*, 169–180.

Topalova, P. (2007). Trade Liberalization, Poverty and Inequality: Evidence from Indian Districts. In *Globalization and Poverty*, NBER Chapters, pp. 291–336. National Bureau of Economic Research, Inc.

Index